

68636

Fr m:

Sent:

To:

Subject:

Li, Ruixiang

Wednesday, June 12, 2002 2:47 PM

STIC-Biotech/ChemLib

Sequence search of Application NO: 09/727,739

Please do a standard search SEQ ID NOs: 3, 15, and 18 against both the commercial and interference amino acid database.

Thank you very much!

Ruixiang Li  
GAU 1646  
CM1 10E18  
Mail Box 10C01  
306-0282

CRFG

Searcher: D. Schreiber  
Phone: 308-4292  
Location: CM1 6A03  
Date Picked Up: 6/13  
Date Completed: 6/19  
Searcher Prep/Review: 7  
Clerical: \_\_\_\_\_  
Online time: 6

## TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: 3  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

## VENDOR/COST(where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: Compugen  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 4.5  
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: June 13, 2002, 12:23:38 : Search time 35.79 Seconds  
(without alignments)  
298.014 Million cell updates/sec

Title: US-09-727-739B-15  
Perfect score: 597  
Sequence: 1 MRVSGIHCALLGLALAIC.....PPRKGAGCKNFYKNGFTSC 111

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR71:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476	79.7	115	2	IS1064
2	194	32.5	125	1	RIAFS2
3	179.5	30.1	116	1	S20630
4	176.5	29.6	116	1	RIHUS1
5	176.5	29.6	116	1	A28968
6	174	29.1	115	2	JC6166
7	163.5	27.4	116	1	RIHUS1
8	163.5	27.4	116	1	RIHUS1
9	162.5	27.2	114	1	RIHUS1
10	156.5	26.2	114	2	IS0798
11	147	24.6	114	2	A20842
12	136	22.8	73	2	S00169
13	136	22.8	92	1	RIHUS1
14	131.5	22.0	121	1	RIAFS1
15	125.5	21.0	74	2	S00166
16	114	19.1	25	2	B60840
17	105	17.6	28	2	A61322
18	104	17.4	34	2	A32271
19	94	15.7	37	2	A32000
20	91.5	15.3	103	2	JC6167
21	87	14.6	112	2	S67489
22	80.5	13.5	105	2	JC5414
23	77.5	13.0	1236	1	VHWVWE
24	77	12.9	14	2	C60414
25	77	12.9	14	2	B60842
26	77	12.9	14	2	B60840
27	77	12.9	14	2	S00172
28	77	12.9	14	2	S00172
29	74	12.4	1241	2	S26373

structural polypro  
structural polypro  
cysteine proteinas  
somatostatin - spo  
ORF2 protein - Orf  
hypothetical prote  
cysteine proteinas  
cyclin g-associate  
F21F23.4 protein -  
subtilisin-like pr  
nitric-oxide synth  
conserved hypotet  
two component sens  
cell cycle histidi  
hypothetical prote  
hypothetical prote

## ALIGNMENTS

### RESULT 1

IS1064

somatostatin II precursor - rainbow trout

C:Species: Oncorhynchus mykiss (rainbow trout)

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999

C:Accession: IS1064

R:Moore, C.A.; Kittilson, J.D.; Dahl, S.K.; Sheridan, M.A.

Gen. Comp. Endocrinol. 98, 253-261, 1995

A:Title: Isolation and characterization of a cDNA encoding for preprosomatostatin coi

A:Reference number: IS1064; MUID:95354921

A:Accession: IS1064

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-115 <MO>

A:Cross-references: EMBL:U32471; NID:975344; PIDN:AC59695.1; PID:975345

C:Superfamily: Somatostatin

Query Match 79.7% Score 476; DB 2; Length 115;

Best Local Similarity 83.1%; Pred. No. 1.4e-43;

Matches 98; Conservative 3; Mismatches 7; Indels 10; Gaps 4;

QY 1 MRVSGIHCALLGLALAICSGAASQPDLDLRSLRLQLRALAALPHRSGVSRWTFY 60

DB 1 MKVCRHICALALLGLALAICSGAASQPDLDLRSLRLQLRALAALPHRSGVSRWTFY 60

QY 61 PNCPCLRWRPRKVKGPOLKAKED---LERSV---DNLPERRKAGCKNFYKNGFTSC 111

DB 61 PNCPCLRWRPRKVKGPOLKAKED---LERSV---DNLPERRKAGCKNFYKNGFTSC 115

RESULT 2

RIAFS2

somatostatin II precursor - American goosfish

C:Species: Lophius americanus (American goosfish)

C>Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 28-May-1999

C:Accession: B93236; A94038; A27376; A01434; A21881; A93236

R:Hobart, P.; Crawford, R.; Shen, L.; Pictet, R.; Rutter, W.J.

Nature 288, 137-141, 1980

A:Title: Cloning and sequence analysis of cDNAs encoding two distinct somatostatin pr

A:Reference number: A93236; MUID:81052423

A:Accession: B93236

A:Molecule type: mRNA

A:Residues: 1-125 <HOB>

A:Cross-references: GB:V00641; GB:J00947; GB:M23199; NID:964030; PIDN:CAA23987.1; PII

A:Experimental source: Islet tissue (endocrine pancreas)

R:Spies, J.; Noe, B.D.

Proc. Natl. Acad. Sci. U.S.A. 82, 277-281, 1985

A:Title: Processing of an anglerfish somatostatin precursor to a hydroxylysine-contai

A:Reference number: A94038; MUID:83113184

A:Accession: A94038

A:Molecule type: protein

A:Residues: 98-125 <SPI>  
 J. Andrews, P.C.; Nichols, R.; Dixon, J.E.  
 J. Biol. Chem. 262, 12692-12699, 1987  
 A:Title: Post-translational processing of preprosomatostatin-II examined using fast atom  
 A:Reference number: A27376; MOID:87308304  
 A:Accession: A27376  
 A:Molecule type: protein  
 A:Residues: 1-76, 'DY', '79-89', 'G', '91-125' <AND>  
 C:Superfamily: somatostatin  
 C:Keywords: hydroxylysine; neuropeptide; pyroglutamic acid  
 F:1-24/Domain: signal sequence #status experimental <SIG>  
 F:25-97/Domain: propeptide #status experimental <PRO>  
 F:97-125/Product: somatostatin II #status experimental <MAT>  
 F:25/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status experimental  
 F:114-125/Disulfide bonds: #status experimental  
 F:120/Modified site: hydroxylysine (Lys) #status experimental

Query Match 32.5%; Score 194; DB 1; Length 125;  
 Best Local Similarity 39.3%; Pred. No. 1.9e-13;  
 Matches 53; Conservative 11; Mismatches 29; Indels 42; Gaps 6;

QY 6 IHICALALLGALALICSGAASQ-----PDLGLARR--LLQALAAALPHRSGVGER 55  
 DB 4 IRCPAILALLLVICGPSVSSQDLREQSDNQLDLRLQHWLLERASAGL-----LSQE 58  
 QY 56 W-----RTFYPCCLWRPRKVKGPOLKAKEDLERSVD---NLPPRR 96  
 DB 59 WSKRVEELIAQMSLPETFO-----READASWATERMNLERSVDNSTNLP 110  
 QY 97 KAGCKNFYWGFTSC 111  
 DB 111 KAGCKNFYWGFTSC 125

RESULT 3  
 somatostatin precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: S20630  
 R:Nata, K.; Kobayashi, T.; Karahashi, K.; Kato, S.; Yamamoto, H.; Yonekura, H.; Okamoto, H.  
 submitted to the EMBL Data Library, June 1991  
 A:Description: Nucleotide sequence determination of chicken somatostatin precursor cDNA  
 A:Reference number: S20630  
 A:Accession: S20630  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <NAT>  
 C:Superfamily: somatostatin

Query Match 30.1%; Score 179.5; DB 1; Length 116;  
 Best Local Similarity 39.8%; Pred. No. 6.2e-12;  
 Matches 49; Conservative 20; Mismatches 27; Indels 27; Gaps 6;

QY 5 QIHICALALLGALALICSGAASQ--SQPDLDLAGRRLLQALAAALPHRSGVGERWRTYFPC 63  
 DB 5 RLQCALALLSTALAVGTVAAPSAPDRL---RQFLQKSLAAA---AGKQELAKYFLAE- 55  
 QY 64 PCLWRPRKVKGPOLKAKE-----DLSRVNLP---PRRKGCKNFYWGKF 108  
 DB 56 --LLSEFQTEALESEDLRGAEQDEVRLELERSANSNPALAPERRKAGCKNFYWGKF 113  
 QY 109 TSC 111  
 DB 114 TSC 116

RESULT 4  
 somatostatin I precursor - human

N:Alternate names: preprosomatostatin  
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 18-Jun-1999  
 C:Accession: A43614; A01430; S09381; S50024  
 R:Shen, L.P.; Rutter, W.J.  
 Science 224, 168-171, 1984  
 A:Title: Sequence of the human somatostatin I gene.  
 A:Reference number: A43614; MOID:84146798  
 A:Accession: A43614  
 A:Molecule type: DNA  
 A:Residues: 1-116 <SH2>  
 A:Cross-references: GB:J00306; NID:g338287; PIDN:AAA60566.1; PID:g338288  
 R:Shen, L.P.; Pictet, R.L.; Rutter, W.J.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 4575-4579, 1982  
 A:Title: Human somatostatin I: sequence of the cDNA.  
 A:Reference number: A01430; MOID:83014931  
 A:Accession: A01430  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <SHE>  
 A:Cross-references: GB:J00306; NID:g338287; PIDN:AAA60566.1; PID:g338288  
 R:Gomez, S.; Boileau, G.; Zollinger, L.; Nault, C.; Rholam, M.; Cohen, P.  
 EMBO J. 8, 2911-2916, 1989  
 A:Title: Site-specific mutagenesis identifies amino acid residues critical in prohom  
 A:Reference number: S09381; MOID:90059875  
 A:Accession: S09381  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 88-108 <GOM>  
 R:Odum, L.; Johnson, A.H.  
 Biochem. J. 303, 263-268, 1994  
 A:Title: Human seminal plasma contains somatostatin-64.  
 A:Reference number: S50024; MOID:95031969  
 A:Accession: S50024  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 53-62; 67-82 <ODU>  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Genetics:  
 A:Gene: GDB:SST  
 A:Cross-references: GDB:119604; OMIM:182450  
 A:Map position: 3q28-3q28  
 A:Introns: 46/3  
 C:Function:  
 A:Description: Inhibits the secretion of a number of peptide hormones, including soma  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; hypothalamus; neuropeptide  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-88/Domain: propeptide #status predicted <PRO>  
 F:89-116/Product: somatostatin-28 #status predicted <M28>  
 F:103-116/Product: somatostatin-14 #status predicted <M14>  
 F:103-116/Disulfide bonds: #status experimental

Query Match 29.6%; Score 176.5; DB 1; Length 116;  
 Best Local Similarity 38.2%; Pred. No. 1.3e-11;  
 Matches 47; Conservative 19; Mismatches 30; Indels 27; Gaps 6;

QY 5 QIHICALALLGALALICSGAASQPDLDLAGRRLLQALAAALPHRSGVGERWRTYFPC 63  
 DB 5 RLQCALALLSTALAVGTVAAPSAPDRL---RQFLQKSLAAA---AGKQELAKYFLAE- 55  
 QY 64 PCLWRPRKVKGPOLKAKE-----DLSRVNLP---PRRKGCKNFYWGKF 108  
 DB 56 --LLSEFQTEALESEDLRGAEQDEVRLELERSANSNPALAPERRKAGCKNFYWGKF 113  
 QY 109 TSC 111  
 DB 114 TSC 116

RESULT 5

A28968  
somatostatin I precursor - crab-eating macaque  
N:Alternate names: preprosomatostatin  
N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
C:Species: Macaca fascicularis (crab-eating macaque)  
C:Date: 30-Jun-1989 #sequence\_revision 31-Jan-1997 #text\_change 18-Jun-1999  
C:Accession: A28968  
R:Travis, G.H.; Sutcliffe, J.G.  
Proc. Natl. Acad. Sci. U.S.A. 85, 1696-1700, 1988  
A:Title: Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low-  
A:Reference number: A28968; MUID:88144503  
A:Accession: A28968  
A:Molecule type: mRNA  
A:Residues: 1-116 <TRA>  
A:Cross-references: GB:M19318; NID:9342298; PIDN:AAA36908.1; PID:9342299  
C:Comment: Somatostatin inhibits the release of somatotropin.  
C:Superfamily: somatostatin  
C:Keywords: hormone; neuropeptide  
F:226Domain: signal sequence #status predicted <SIG>  
F:25-86/Domain: propeptide #status predicted <PRO>  
F:89-116/Product: somatostatin-28 #status predicted <M28>  
F:103-116/Product: somatostatin-14 #status predicted <M14>  
F:105-116/Disulfide bonds: #status predicted

Query Match 29.6%; Score 176.5; DB 1; Length 116;  
Best Local Similarity 38.2%; Pred. No. 1.3e-11;  
Matches 47; Conservative 19; Mismatches 30; Indels 27; Gaps 6;  
6;  
QY 5 QIHCALALGLALAI-CSQGAASOPDLDLASRLIQRALAAALPHRSGVSEWRTFFPNC 63  
DB 5 RIQCALAALSIVLALGCVTGAAPDRL----RQFLQKSLAA-----AGQELAKYFLAE- 55  
QY 64 PCRWRPRVKPGPKAKE-----DLSRVNDLP---PRRKAGCKNFKYKGF 108  
DB 56 --LLSEPNGTENDALEPEDLSQAABQDEMRLELQSGANSNPAMAPRRKAGCKNFKWTF 113  
QY 109 TSC 111  
DB 114 TSC 116

RESULT 6  
JC6166  
somatostatin-14 precursor - laughing frog  
N:Alternate names: PSS1 protein  
C:Species: Rana ridibunda (laughing frog)  
C:Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 16-Jul-1999  
C:Accession: JC6166  
R:Tostivint, H.; Lihmann, I.; Buchares, C.; Vieau, D.; Coulouarn, Y.; Fournier, A.; Co  
Proc. Natl. Acad. Sci. U.S.A. 93, 12605-12610, 1996  
A:Title: Occurrence of two somatostatin variants in the frog brain: Characterization of  
A:Reference number: JC6166; MUID:97057290  
A:Contents: brain  
A:Accession: JC6166  
A:Molecule type: mRNA  
A:Residues: 1-115 <TOS>  
A:Cross-references: GB:U68136; NID:gl890650; PIDN:AC60093.1; PID:gl890651  
C:Comment: This protein acts both as a neurotransmitter/neuromodulator and a hormone.  
C:Gene: pss1  
C:Superfamily: somatostatin  
C:Keywords: brain; hormone

Query Match 29.1%; Score 174; DB 2; Length 115;  
Best Local Similarity 39.4%; Pred. No. 2.4e-11;  
Matches 50; Conservative 17; Mismatches 32; Indels 28; Gaps 6;  
6;  
QY 1 MRYSOIHCALILGLALICSGAA-SOPDLASRLIQRALAAALPHRSGVSEWRTFF 59  
DB 1 MQSCRVOCALTLLSLALAINISIAPTDPL-----RQFLQKSLAA-----GKELAKYF 51

QY 60 YPNCPCLRWPRKVRKVPOLKAKE-----DLSRVNDLP---PRRKAGCKNFKY 104  
DB 52 LAE---LLSEPSQTDNEALESDDLPRGAEQDEVLELERSANSALAPRRKAGCKNFK 108  
QY 105 WKGTFTSC 111  
DB 109 WKGTFTSC 115

RESULT 7  
RIRTS1  
somatostatin precursor - rat  
N:Alternate names: preprosomatostatin  
N:Contains: somatostatin-14; somatostatin-28  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 03-Aug-1984 #sequence\_revision 02-Aug-1984 #text\_change 18-Jun-1999  
C:Accession: A20983; A01431; A47598; A22529; I55220; I51829  
R:Montminy, M.R.; Goodman, R.H.; Horovitch, S.J.; Habener, J.F.  
Proc. Natl. Acad. Sci. U.S.A. 81, 3337-3340, 1984  
A:Title: Primary structure of the gene encoding rat preprosomatostatin.  
A:Reference number: A20983; MUID:84221954  
A:Accession: A20983  
A:Molecule type: DNA  
A:Residues: 1-116 <NON>  
A:Cross-references: GB:J00787; NID:g207024; PIDN:AAA42164.1; PID:g207025  
A:Note: The authors translated the codon ACC for residue 43 as Tyr  
R:Argos, P.; Taylor, W.L.; Minth, C.D.; Dixon, J.E.  
J. Biol. Chem. 258, 8788-8793, 1983  
A:Title: Nucleotide and amino acid sequence comparisons of preprosomatostatins.  
A:Reference number: A01431; MUID:83238516  
A:Accession: A01431  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-116 <ARG>  
R:Benoit, R.; Ling, N.; Esch, F.  
Science 238, 1128-1129, 1987  
A:Title: A new preprosomatostatin-derived peptide reveals a pattern for prohormone clea  
A:Reference number: A47598; MUID:88070564  
A:Accession: A47598  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 25-34 <BEN>  
R:Tavianini, M.A.; Hayes, T.E.; Magazini, M.D.; Minth, C.D.; Dixon, J.E.  
J. Biol. Chem. 259, 11798-11803, 1984  
A:Title: Isolation, characterization, and DNA sequence of the rat somatostatin gene.  
A:Reference number: A22529; MUID:85006903  
A:Accession: A22529  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-116 <NAV>  
A:Cross-references: GB:K02248; NID:g207014; PIDN:AAA42161.1; PID:g207017  
R:Goodman, R.H.; Jacobs, J.W.; Dee, P.C.; Habener, J.F.  
J. Biol. Chem. 257, 1156-1159, 1982  
A:Title: Somatostatin-28 encoded in a cloned cDNA obtained from a rat medullary thyr  
A:Reference number: I55220; MUID:82120034  
A:Accession: I55220  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 38-78, H', 80-116 <RES>  
A:Cross-references: GB:J00788; NID:g207018; PIDN:AAA42162.1; PID:g207019  
R:Goodman, R.H.; Montminy, M.R.; Low, M.J.; Habener, J.F.  
Adv. Exp. Med. Biol. 188, 31-47, 1985  
A:Title: Biosynthesis of rat preprosomatostatin.  
A:Reference number: I51829; MUID:85303584  
A:Accession: I51829  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-116 <RES>  
A:Cross-references: GB:M25890; NID:g207030; PIDN:AAA42167.1; PID:g207031  
C:Comment: Somatostatins are found in a variety of tissues, including hypothalamus,  
C:Genetics: 46/3  
A:Introns: 46/3  
C:Function:

A:Description: inhibits the secretion of a number of peptide hormones, including somatostatin  
C:Superfamily: somatostatin  
C:Keywords: hormone; hypothalamus; neuropeptide  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-88/Domain: propeptide #status experimental <PRO>  
F:89-116/Product: somatostatin-28 #status predicted <S28>  
F:103-116/Product: somatostatin-14 #status predicted <M14>  
F:105-116/Disulfide bonds: #status predicted

Query Match 27.4%; Score 163.5; DB 1; Length 116;  
Best Local Similarity 37.4%; Pred. No. 3.1e-10;  
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;  
QY 5 QTHCALALGLALAICS-OGAASQPDLDLASRRLLQALAAALPHRSVGSERWTFYPNC 63  
DB 5 RLQCALAALCIVLAGVGTGAPSDRL-----RQFLQKSLAA-----TKQELAKYFLAE- 55  
QY 64 PCLWRPRKVKGPOLKAKE-----DLERSVDNLP---PRRKAGCKNFYKGF 108  
DB 56 --LLSEPNQTDALPEPDLQAAEQDEMRLQRSANSNPAMAPRRKAGCKNFYKTF 113  
QY 109 TSC 111  
DB 114 TSC 116

RESULT 8  
RIMSS1  
somatostatin precursor - mouse  
N:Alternate names: preprosomatostatin  
N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Jan-1997 #text\_change 18-Jun-1999  
C:Accession: S08416  
R:Fuhrmann, G.; Heilig, R.; Kempf, J.; Ebel, A.  
Nucleic Acids Res. 18, 1287, 1990.  
A:Title: Nucleotide sequence of the mouse preprosomatostatin gene.  
A:Reference number: S08416; MUID:90206793  
A:Molecule type: DNA  
A:Accession: S08416  
A:Residues: 1-116 <FUH>  
A:Cross-references: EMBL:X51468; NID:g53768; PIDN:CAA35831.1; PID:g297530  
C:Comment: Somatostatin inhibits the release of somatotropin.  
C:Genetics:  
A:Introns: 46/3  
C:Superfamily: somatostatin  
C:Keywords: hormone; neuropeptide  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-88/Domain: propeptide #status predicted <PRO>  
F:89-116/Product: somatostatin-28 #status predicted <M28>  
F:103-116/Product: somatostatin-14 #status predicted <M14>  
F:105-116/Disulfide bonds: #status predicted

Query Match 27.4%; Score 163.5; DB 1; Length 116;  
Best Local Similarity 37.4%; Pred. No. 3.1e-10;  
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;  
QY 5 QTHCALALGLALAICS-OGAASQPDLDLASRRLLQALAAALPHRSVGSERWTFYPNC 63  
DB 5 RLQCALAALCIVLAGVGTGAPSDRL-----RQFLQKSLAA-----TKQELAKYFLAE- 55  
QY 64 PCLWRPRKVKGPOLKAKE-----DLERSVDNLP---PRRKAGCKNFYKGF 108  
DB 56 --LLSEPNQTDALPEPDLQAAEQDEMRLQRSANSNPAMAPRRKAGCKNFYKTF 113  
QY 109 TSC 111  
DB 114 TSC 116

RESULT 9

RIBOS1  
somatostatin precursor - bovine  
N:Alternate names: preprosomatostatin  
N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 06-Mar-1992 #sequence\_revision 31-Jan-1997 #text\_change 18-Jun-1999  
C:Accession: A40929  
R:Su, C.J.; White, J.W.; Li, W.H.; Luo, C.C.; Frazier, M.L.; Saunders, G.F.; Chan, L.  
Mol. Endocrinol. 2, 209-216, 1988  
A:Title: Structure and evolution of somatostatin genes.  
A:Reference number: A40929; MUID:88288237  
A:Accession: A40929  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-116 <SU>  
A:Cross-references: GB:M31217; NID:g163636; PIDN:AAA30744.1; PID:g163637  
A:Note: the authors translated the codon ATT for residue 65 as Asn  
C:Comment: Somatostatin inhibits the release of somatotropin.  
C:Superfamily: somatostatin  
C:Keywords: hormone; neuropeptide  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-88/Domain: propeptide #status predicted <PRO>  
F:89-116/Product: somatostatin-28 #status predicted <M28>  
F:103-116/Product: somatostatin-14 #status predicted <M14>  
F:105-116/Disulfide bonds: #status predicted

Query Match 27.2%; Score 162.5; DB 1; Length 116;  
Best Local Similarity 37.4%; Pred. No. 4e-10;  
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;  
QY 5 QTHCALALGLALAICS-OGAASQPDLDLASRRLLQALAAALPHRSVGSERWTFYPNC 63  
DB 5 RLQCALAALCIVLAGVGTGAPSDRL-----RQFLQKSLAA-----AGQELAKYFLAE- 55  
QY 64 PCLWRPRKVKGPOLKAKE-----DLERSVDNLP---PRRKAGCKNFYKGF 108  
DB 56 --LLSEPNQTDALPEPDLQAAEQDEMRLQRSANSNPAMAPRRKAGCKNFYKTF 113  
QY 109 TSC 111  
DB 114 TSC 116

RESULT 10  
RIBDS1  
somatostatin-14 precursor - channel catfish  
N:Alternate names: somatostatin I  
N:Contains: somatostatin-14  
C:Species: Ictalurus punctatus (channel catfish)  
C>Date: 30-Jun-1980 #sequence\_revision 31-Dec-1993 #text\_change 18-Jun-1999  
C:Accession: S00292; A93897; A92334; A01435  
R:Mintz, C.D.; Taylor, W.L.; Magazin, M.; Tavlanini, M.A.; Collier, K.; Weith, H.L.;  
J. Biol. Chem. 257, 10372-10377, 1982  
A:Title: The structure of cloned DNA complementary to catfish pancreatic somatostatin  
A:Reference number: S00292; MUID:82265698  
A:Accession: S00292  
A:Molecule type: mRNA  
A:Residues: 1-114 <MIN>  
A:Cross-references: EMBL:V00607; NID:g64017; PIDN:CAA23877.1; PID:g64018  
R:Taylor, W.L.; Collier, K.J.; Deschenes, R.J.; Weith, H.L.; Dixon, J.E.  
Proc. Natl. Acad. Sci. U.S.A. 78, 6694-6698, 1981  
A:Title: Sequence analysis of a cDNA coding for a pancreatic precursor to somatostatin  
A:Reference number: A93897; MUID:82082515  
A:Accession: A93897  
A:Molecule type: mRNA  
A:Residues: 82-108 <TAY>  
A:Cross-references: GB:J00944  
R:Andrews, P.C.; Dixon, J.E.  
J. Biol. Chem. 256, 8267-8270, 1981  
A:Title: Isolation and structure of a peptide hormone predicted from a mRNA sequence  
A:Reference number: A92334; MUID:81264223  
A:Accession: A92334



A.Molecule type: Protein  
A.Residues: 1-92 <SC3>  
A.Experimental source: Intestine  
R.P:Reddall, A.V.; Huang, W.Y.; Chang, R.C.C.; Arimura, A.; Redding, T.W.; Millar, R.P.; H  
FEBs Lett. 109, 5336, 1980  
A.Title: N-terminally extended somatostatin: the primary structure of somatostatin-28.  
A.Reference number: A91273; MUID:80113258  
A.Accession: A91273  
A.Molecule type: protein  
A.Residues: 65-92 <PRA>  
A.Experimental source: Intestine  
R.P:Reddall, A.V.; Huang, W.Y.; Chang, R.C.C.; Arimura, A.; Redding, T.W.; Millar, R.P.; H  
Proc.Natl. Acad. Sci. U.S.A. 77, 4489-4493, 1980  
A.Title: Isolation and structure of pro-somatostatin: a putative somatostatin precursor  
A.Reference number: A93854; MUID:81054799  
A.Accession: A93854  
A.Molecule type: protein  
A.Residues: 65-92 <SCH>  
A.Experimental source: Hypothalamus  
R.P:Reddall, A.V.; Dupont, A.; Arimura, A.; Redding, T.W.; Nishi, N.; Linthicum, G.L.; Sch  
Biochemistry 15, 509-514, 1976  
A.Title: Isolation and structure of somatostatin from porcine hypothalamus.  
A.Reference number: A90398; MUID:76136331  
A.Accession: A90398  
A.Molecule type: protein  
A.Residues: 79-92 <SC2>  
A.Experimental source: Hypothalamus  
R:Bersani, M.; Johnsen, A.H.; Holst, J.J.  
FEBs Lett. 279, 237-239, 1991  
A.Title: Oxidation/reduction explains heterogeneity of pancreatic somatostatin.  
A.Reference number: S13616; MUID:91160722  
A.Accession: S13616  
A.Molecule type: protein  
A.Residues: 79-92 <BE2>  
C:Comment: Somatostatin inhibits the release of somatotropin.  
C:Superfamily: somatostatin  
C:Keywords: hormone; hypothalamus; intestine; neuropeptide  
F:1-64/Domain: propeptide #status experimental <PRO>  
F:65-92/Product: somatostatin-28 #status experimental <M28>  
F:79-92/Product: somatostatin-14 #status experimental <M14>  
F:79-92/Binding site: carbohydrate (asn) (covalent) #status absent  
F:81-92/Disulfide bonds: #status experimental

Query Match	22.8%;	Score 136;	DB 1;	Length 92;
Best Local Similarity	36.9%;	Pred. No. 2.1e-07;		
Matches	38;	Conservative 15;	Mismatches 24;	Indels 26; Gaps 5;

24 AASOPDLDLASRRLQORALAAALPHRSVGVSERWTFYPNCPLRWPRKVKGPQLKAKE- 82  
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
1 APSDPRL-----ROFLOKSLAAA-----AGKOELAKYFLAE-----LLSEPNOTENDALEPDL 49

```

83 -----DLERSVDNLP--PRERKAGCKNFYWKGF TSC 111
      :|:| :| | | | | | | | | | | | | | |
50 SQAAEQDEMRLEORSANSNPANAPRERKAGCKNFWKFT TSC 92

```

**RESULT . 15**

RIAFSI

RIAFSI  
somatostatin I precursor - American goosefish

**N; Contains: somatostatin I**

C; Species: *Lophius americanus* (American goosefish)

C; Species: *Lophyrus americanus* (American goosetern)  
C; Date: 31-Mar-1981 #sequence\_revision 31-Mar-1981 #text\_change 28-May-1999

C;Accession: A93236; A93860; A91087; A01433

R;Hobart, P.; Crawford, R.; Shen, L.; Pictet, R.; Rutter, W.J.

K, HODDGE, F., CLAWSON, K.  
Nature 288, 137-141, 1980

**A:Title: Cloning and sequence analysis of**

A:Reference number: A9

A; reference number:  
A: Accession: A93236

A; ACCESSION: A93236  
A: MOLECULE TYPE: mRNA

A; Molecule type: mRNA  
A: Residues: 1-121 <HOP

A;Residues: I-IZI <HOB>  
A:Cross-references: GB:V00640: GB:J00946: NID:q64028: PIDN:CAA23986.1: PID:q64029

Cross-References: GB:V00840; GB:J00946; NID:964028; FIDN:CAAZ3988.1; FID:964028; R:Goodman, R.H.; Jacobs, J.W.; Chln. W.W.; Lund, P.K.; Dee, P.C.; Habener, J.F.

R; GOODMAN, R.H.; JACOBS, J.W.; CHIN, W.W.; LUDU, Proc. Natl. Acad. Sci. U.S.A. 77. 5869-5873. 1980

A>Title: Nucleotide sequence of a cloned structural gene coding for a precursor of  
A.Reference number: A93860; PMID:81077276  
A.Accession: A93860  
A.Modifier: 17-20, 'V', 22-82, 'E', 84-121 <GOO>  
A.Experimental source: islet tissue (endocrine pancreas)  
R.German, R. H. Jacobs, J. W. Chin, W. W. Lund, P. K.; Dee, P. C.; Habener, J. F.  
Proc Natl Acad Sci U S A. 79, 1682, 1982  
A.Reference number: A93905  
A.Contents: annotation; erratum  
R.Noe, B. D.; Spiess, J.; Rivier, J. E.; Vale, W.  
Endocrinology 105, 1410-1415, 1979  
A>Title: Isolation and characterization of somatostatin from anglerfish pancreatic i  
A.Reference number: A91087; PMID:80046482  
A.Accession: A91087  
A.Molecule type: protein  
A.Residues: 108-121 <NOE>  
C:Superfamily: somatostatin  
C:Keywords: neuropeptide  
F:1-24/Domain: signal sequence \$status predicted <SIG>  
F:108-121/Product: somatostatin I \$status experimental <MAT>  
F:110-121/Disulfide bonds: \$status predicted

```
Query Match      22.08; Score 131.5; DB.1; Length 121;  
Best Local Similarity 37.9%; Fred.No. 8.5e-07;  
Matches 47; Conservative 5; Mismatches 47; Indels 25; Gaps  
  
    4 SQTHCALLLGLALIAICSGGAASOPDLDIASRRLLQR-----ALAALPHRS 50  
       !::!::::!!:  
    7 SRLRCLVVLLLSTASTSCSPAGORSKL--RLLRHYPLQGGKODMTRSLAELLSDL 64  
       !!:!!!  
  
   51 GVSEWRRTYFNPCCILRWPRPKVGPQLAKEDLEERSVDN---LPPPERKAGCKNFYWKG 107  
       !!!:!!!!!  
   65 LOGENALEEENEFPFAE-----GGPE-DAAHDLERAAAGGPGLLAPRKERAGCKNEFWKT 117  
  
  108 FTSC 111  
     !!!!!  
  118 FTSC 121
```

Search completed: June 13, 2002, 12:23:38  
Job time: 196 sec

Fri Jan 14 08:03:45 2002

us-09-727-739b-15.rpr

Page 7





```

RESULT 2
Q90Y43 ID Q90Y43 PRELIMINARY; PRT; 115 AA.
AC Q90Y43; 01-DEC-2001 (TREMELREL. 19, Created)
DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMELREL. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Osteoglossum bicirrhosum (silver arowana).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxID=109271;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292650; AAK97067.1;
SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

Query Match 39.5%; Score 236; DB 13; Length 115;
Best Local Similarity 47.5%; Pred. No. 9.3e-19;
Matches 58; Conservative 10; Mismatches 36; Indels 18; Gaps 4;

QY 1 MRVSOIHCALALGLAICSGAASQPDLDLASRLRLQALAAALPHRSGVSRWTFY 60
DB 1 MKICQIHCTLVLLGLVGLYCPSAASQPDLDLASRLRLQALAAALPHRSGVSRWTFY 60

QY 61 PNCPCLEWRPRKVKQ-----LKAKEDLERSVD--NLPPRKGCKNFYWKGT 109
DB 54 VEELSLRLAPAGEVPGAVSADEEDVRVLDLERSLELNLPRLPRKAGCKNFYWKGT 113

QY 110 SC 111
DB 114 SC 115

RESULT 3
Q90Y42 ID Q90Y42 PRELIMINARY; PRT; 114 AA.
AC Q90Y42; 01-DEC-2001 (TREMELREL. 19, Created)
DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMELREL. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292651; AAK97068.1;
SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 38.3%; Score 228.5; DB 13; Length 114;
Best Local Similarity 41.4%; Pred. No. 6.3e-18;
Matches 55; Conservative 14; Mismatches 23; Indels 41; Gaps 5;

QY 1 MRVSOIHCALALGLAICSGAASQPDLDLASRLRLQALAAALPHRSGVSRWTFY 60
DB 1 MKICQIHCTLVLLGLVGLVCGSSSATO--LDSRYSLVQARAASMGPD----- 48

```

```

QY 61 PNCPCLEWRPRKVK-----GPKLKAKE-----DLERSVD--NLPPRKA 98
DB 49 -----WGKLSVEDLSLAATEADNPFQDMSAAEESGAILDLERSVEGPNVPRKA 101

QY 99 GCKNFYWKGTSC 111
DB 102 GCKNFYWKGTSC 114

RESULT 4
Q90Y41 ID Q90Y41 PRELIMINARY; PRT; 114 AA.
AC Q90Y41; 01-DEC-2001 (TREMELREL. 19, Created)
DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMELREL. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Gnathonemus petersi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Mormyridae; Gnathonemus.
OX NCBI_TaxID=42645;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292652; AAK97069.1;
SQ SEQUENCE 114 AA; 12494 MW; 454DA57A309CA8F2 CRC64;

Query Match 32.2%; Score 192.5; DB 13; Length 114;
Best Local Similarity 41.4%; Pred. No. 6.5e-14;
Matches 55; Conservative 8; Mismatches 29; Indels 41; Gaps 5;

QY 1 MRVSOIHCALALGLAICSGAASQPDLDLASRLRLQALAAAL-----PHRSGV 52
DB 1 MLSRIQCALALLSLALAVSVSAAPS---DLKRLQLORSLLAPASKQDLARNPLELL 57

QY 53 SERWTFYPCPLWRPRKVKGPOLKAKEDLERSVDN-----LPPRKA 98
DB 58 SEMVR-----VENEALP-DDLSRGADQEVRELERAAGPALAPRKA 101

QY 99 GCKNFYWKGTSC 111
DB 102 GCKNFYWKGTSC 114

RESULT 5
Q90Y40 ID Q90Y40 PRELIMINARY; PRT; 114 AA.
AC Q90Y40; 01-DEC-2001 (TREMELREL. 19, Created)
DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMELREL. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Chitala chitala.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Notopteridae; Chitala.
OX NCBI_TaxID=112163;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292653; AAK97070.1;
SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F58E34F971 CRC64;

```





ID Q9P2X3 PRELIMINARY; PRT: 1242 AA.  
 AC Q9P2X3; 2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE STRUCTURAL POLYPEPTIDE.  
 OS Eastern equine encephalitis virus (Eastern equine encephalomyelitis virus).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;  
 OC Alphavirus.  
 OX NCBI\_TaxID=11021;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PL93-939;  
 RA Brault A.C., Powers A.M., Kang W., Tesh R.B., Shope R.E., Weaver S.C.;  
 RT "Genetic and Antigenic Diversity among Eastern Equine encephalitis  
 RL viruses from North Central and South America.";  
 DR Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF159554; AAF04796.1;  
 DR HSP; P03315; IYCP.  
 DR InterPro: IPR001836; Alpha\_core;  
 DR InterPro: IPR002548; Alpha\_E1\_glycop.  
 DR InterPro: IPR000936; Alpha\_E2\_glycop.  
 DR InterPro: IPR002533; Alpha\_E3\_glycop.  
 DR InterPro: IPR000930; Togavirin.  
 DR Pfam: PF00944; Alpha\_core; 1.  
 DR Pfam: PF01589; Alpha\_E1\_glycop; 1.  
 DR Pfam: PF00943; Alpha\_E2\_glycop; 1.  
 DR Pfam: PF01563; Alpha\_E3\_glycop; 1.  
 DR PRINTS: PR00798; TOGAVIRIN.  
 KW Polyprotein.  
 SQ SEQUENCE 1242 AA; 137613 MW; 6F474E82A91FF4CD CRC64;

Query Match 12.9%; Score 77; DB 12; Length 1242;  
 Best Local Similarity 44.0%; Pred. No. 6.8;  
 Matches 22; Conservative 3; Mismatches 9; Indels 16; Gaps 3;  
 QY 61 PNCPCLRWRPRKVKGPOLKAK-EDLERSVDNL-----PPPERK 97  
 DB 23 PNPFRWRPRK---PPLAQIEDLRSLANLTQKRAPNPAGPPAKRK 69

RESULT 14  
 ID Q98HG9 PRELIMINARY; PRT: 251 AA.  
 AC Q98HG9;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DE MLL2874. PROTEIN.  
 DE MLL2874.  
 GN MLL2874.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RL Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL: AP003000; BAB49897.1;  
 KW Complete proteome.  
 SQ SEQUENCE 251 AA; 27987 MW; 45709C6A71EDD2AA CRC64;

Query Match 12.7%; Score 76; DB 16; Length 251;  
 Best Local Similarity 33.3%; Pred. No. 1.5;  
 Matches 29; Conservative 15; Mismatches 23; Indels 20; Gaps 5;  
 QY 3 VSGIHCALALGLALAIQSGAASQPDLDLASRLQLRALAALPHRSVSRWTFYPN 62  
 DB 180 VSAIKTRGGLALKIA---RVHARGDLIAARFL---AVAAAPKRR--SEAWR----- 226  
 QY 63 CPCLWR-----PRKVKGPOLKAKEDL 84  
 DB 227 --CLAYRLKLAVRRLSAPKIELQSVL 251

RESULT 15  
 Q86972 PRELIMINARY; PRT: 259 AA.  
 ID Q86972  
 AC Q86972;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NUCLEOCAPSID GENE C (FRAGMENT).  
 OS Western equine encephalitis virus.  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;  
 OC Alphavirus.  
 OX NCBI\_TaxID=11039;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96097271; PubMed=8535272;  
 RA Uryvaev L.V., Lebedev A.Iu., Sokolova T.M., Iuferov V.P.;  
 RT "[Primary structure of the nucleocapsid gene C and the protein coded  
 RL by it from the Western equine encephalomyelitis virus].";  
 DR EMBL: S80862; AAB35859.1;  
 DR HSP; P03315; IYCP.  
 DR InterPro: IPR001836; Alpha\_core.  
 DR InterPro: IPR000930; Togavirin.  
 DR Pfam: PF00944; Alpha\_core; 1.  
 DR PRINTS: PR00798; TOGAVIRIN.  
 FT NON\_TER 259  
 SQ SEQUENCE 259 AA; 29225 MW; 8989AF2668A25CE CRC64;

Query Match 12.6%; Score 75.5; DB 12; Length 259;  
 Best Local Similarity 38.5%; Pred. No. 1.8;  
 Matches 20; Conservative 7; Mismatches 8; Indels 17; Gaps 3;  
 QY 61 PNCPCLRWRPRKVKGPOLKAK-EDLERSVDNL-----PPPERKA 98  
 DB 23 PNPFRWRPRK---PPLAQIEDLRSLANLTQKRAPNPAGPPAKRK 71

Search completed: June 13, 2002, 12:31:02  
 Job time: 484 sec





FT Peptide 98..111  
/note= "SS-14 variant peptide"  
XX CA2325169-A1.  
XX 03-JUN-2001.  
XX 01-DEC-2000; 2000CA-2325169.  
XX 03-DEC-1999; 99US-0168934.  
XX (NDSU-) NDSU RES FOUND.  
XX Sheridan MA, Moore CA, Kittelson JD;  
XX WPI; 2001-425997/46.  
XX N-PSDB; AAS12935.  
XX  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
XX immunodeficiency syndrome and neurological disorders  
XX  
XX Claim 1; Fig 3; 52pp; English.  
XX  
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
XX preprosomatostatin II (PPSS-II). The protein sequences and their  
XX associated polynucleotides are useful for identifying modified  
XX somatostatin polypeptides which functions as a somatostatin agonist useful  
XX for research, therapeutics or diagnostics, including medical and  
XX veterinary applications. The wild-type somatostatin and its modified  
XX version are useful for treating hypersecretion from endocrine tumours in  
XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
XX through their effects on cell proliferation and apoptosis and as adjuncts  
XX in the treatment of diabetes mellitus via inhibition of growth hormone  
XX and glucagon. In addition, dysfunctional somatostatin secretion is  
XX associated with acquired immunodeficiency syndrome (AIDS) and various  
XX neurological disorders (e.g. epilepsy, Alzheimer's disease and  
XX Huntington's disease) and somatostatin antagonists are effective in the  
XX treatment of such conditions. Nucleic acids encoding the polypeptides are  
XX useful in gene therapy and fusion peptides can be targeted to neoplasms  
XX and their metastases, inhibiting the release of their secretory products.  
XX This sequence represents O. Mykiss PPSS-II', protein.  
XX Note: The features for this sequence are specifically claimed in the  
XX specification.

Sequence 111 AA;

Query Match 100.0%; Score 597; DB 22; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1.6e-63;  
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRVSIHICALALLGLAICSGQAQSPDLASRLRLQALAAALPHRSGVSRWTFY 60  
Db 1 mrvsgihcalallglalaicsgqaasqpdldasrrllqralaaalphrsgvserwtfy 60  
QY 61 PNCPCPLRWRPRKVPQPKAKEDLERSVDNLPRLPRKAGCKNFYWKGFSC 111  
Db 61 pncpcplrwrprkvkqpkakadlersvdnlprerkagcknfykqgfsc 111

RESULT 2  
AAU07667  
ID AAU07667 standard; Protein; 115 AA.  
XX  
XX AAU07667;  
XX  
XX 04-DEC-2001 (first entry)  
XX  
XX Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.  
XX

KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.  
XX  
XX Oncorhynchus mykiss.  
XX  
XX Location/Qualifiers  
XX Key 1..25  
XX Peptide /note= "Signal peptide"  
XX Protein 1..87  
XX /note= "PPSS-II' pre-sequence"  
XX Protein 26..115  
XX /note= "Mature PPSS-II"  
XX Misc-difference 74  
XX /note= "Encoded by CAA"  
XX Peptide 88..101  
XX /note= "PPSS-II' pro-sequence"  
XX Peptide 88..115  
XX /note= "Prosomatostatin II"  
XX Cleavage-site 100..101  
XX /note= "Dibasic cleavage site"  
XX Peptide 102..115  
XX /note= "SS-14 variant peptide"  
XX  
XX CA2325169-A1.  
XX  
XX 03-JUN-2001.  
XX  
XX 01-DEC-2000; 2000CA-2325169.  
XX  
XX 03-DEC-1999; 99US-0168934.  
XX (NDSU-) NDSU RES FOUND.  
XX  
XX Sheridan MA, Moore CA, Kittelson JD;  
XX WPI; 2001-425997/46.  
XX N-PSDB; AAS12934.  
XX  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
XX immunodeficiency syndrome and neurological disorders  
XX  
XX Claim 2; Fig 3; 52pp; English.  
XX  
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
XX preprosomatostatin II (PPSS-II). The protein sequences and their  
XX associated polynucleotides are useful for identifying modified  
XX somatostatin polypeptides which functions as a somatostatin agonist useful  
XX for research, therapeutics or diagnostics, including medical and  
XX veterinary applications. The wild-type somatostatin and its modified  
XX version are useful for treating hypersecretion from endocrine tumours in  
XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
XX through their effects on cell proliferation and apoptosis and as adjuncts  
XX in the treatment of diabetes mellitus via inhibition of growth hormone  
XX and glucagon. In addition, dysfunctional somatostatin secretion is  
XX associated with acquired immunodeficiency syndrome (AIDS) and various  
XX neurological disorders (e.g. epilepsy, Alzheimer's disease and  
XX Huntington's disease) and somatostatin antagonists are effective in the  
XX treatment of such conditions. Nucleic acids encoding the polypeptides are  
XX useful in gene therapy and fusion peptides can be targeted to neoplasms  
XX and their metastases, inhibiting the release of their secretory products.  
XX This sequence represents O. Mykiss PPSS-II' protein.  
XX Note: The features for this sequence are specifically claimed in the  
XX specification.



SQ Sequence 115 AA;  
 Query Match 80.7%; Score 482; DB 22; Length 115;  
 Best Local Similarity 83.9%; Pred. No. 9.3e-50;  
 Matches 99; Conservative 3; Mismatches 6; Indels 10; Gaps 4;  
 QY 1 MRVSIHCALALLGLAICSGAASQPDLDLASRLQALAAALPHRSVGSERWTFY 60  
 Db 1 mkvcrihcalallglalaicsgaasqpdldlrslrrllqraaaalphrsvgservtfy 60  
 QY 61 PNCPCLRWRPRKVGKPOLKAKED----LERSV---DNLPPRRKAGCKNFYKGFSTC 111  
 Db 61 pncpcl--rprkvkcp-agakeldlrvelersvgnpnlprrerkgcknfykgtsc 115  
 RESULT 3  
 AAP20029  
 ID AAP20029 standard; Protein; 125 AA.  
 XX AAP20029;  
 DT 14-AUG-1992 (first entry)  
 DE Sequence of preprosomatostatin-2 encoded on pLas2.  
 XX Somatostatin; growth hormone; peptide hormone; secretion.  
 KW Key Location/Qualifiers  
 FH 112..125  
 FT /label= Somatostatin II  
 PN EP46669-A.  
 XX 03-MAR-1982.  
 XX 21-AUG-1981; 81EP-0303825.  
 XX 25-AUG-1980; 80US-0181046.  
 XX (REGC ) UNIV OF CALIFORNIA.  
 XX Hobart P, Crawford R, Pictet RL, Rutter WJ;  
 DR WPI; 1982-18113E/10.  
 DR N-PSDB; AAN20034.  
 XX New somatostatin and precursors - produced by transformed  
 PT microorganisms  
 PS Example; Fig 3; 50pp; English.  
 CC The inventors claim preprosomatostatin-1, prosomatostatin-1,  
 CC preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA  
 CC encoding them. The translation of somatostatin mRNA yields a  
 CC precursor (propro S1) containing a signal peptide which may be  
 CC released during the transit into the endoplasmic reticulum, and the  
 CC resultant precursor (pro S1) is subsequently cleaved to yield S1  
 CC itself. The prepeptide portion of propro S1 is probably about 20-25  
 CC bases long. Translation of pLas2 predicts the sequence of a 125 AA  
 CC peptide which surprisingly contains a 14 AA sequence at its carboxy  
 CC terminus which differs from S1 by only 2 AAs, and is termed  
 CC Somatostatin 2 (S2).  
 XX SQ Sequence 125 AA;

Query Match 32.5%; Score 194; DB 3; Length 125;  
 Best Local Similarity 39.3%; Pred. No. 2.8e-15;  
 Matches 53; Conservative 11; Mismatches 29; Indels 42; Gaps 6;  
 QY 6 IHCALALLGLAICSGAASQ-----PDLASRR--LLQALAAALPHRSVGSER 55  
 Db 6 ihcacalallglalaicsgaasq-----pdlasrr--llqalaaalphrsvgserv 55

Db 4 ircpallallvicgspvsqldreqsdnqldldlrlqhwllersag1-----lsqe 58  
 QY 56 W-----RTFYPCPCLRWRPRKVGKPOLKAKEDLERSVD---NLPPRR 96  
 Db 59 wkraveellagmalpeatf-----reaedasmategmrlersvdstnnlprrer 110  
 QY 97 KAGCKNFYKGFSTC 111  
 Db 111 kagcknfykgtsc 125  
 RESULT 4  
 AAG03774  
 ID AAG03774 standard; Protein; 116 AA.  
 XX AAG03774;  
 XX 06-OCT-2000 (first entry)  
 DT Human secreted protein, SEQ ID NO: 7855.  
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.  
 XX Homo sapiens.  
 OS EP1033401-A2.  
 PN 06-SEP-2000.  
 PD 21-FEB-2000; 2000EP-0200610.  
 PF 26-FEB-1999; 99US-0122487.  
 PR (GEST ) GENSET.  
 XX Dumas-Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR N-PSDB; AAC03780.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 13; SEQ ID 7855; 71pp + CD-ROM; English.  
 XX The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dr primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 XX SQ Sequence 116 AA;

Query Match 29.6%; Score 176.5; DB 21; Length 116;  
 Best Local Similarity 38.2%; Pred. No. 5.2e-13;  
 Matches 47; Conservative 19; Mismatches 30; Indels 27; Gaps 6;  
 QY 5 QIHCALALLGLALAI-CSQGAASQPDLDLASRLQALAAALPHRSVGSERWTFY 63  
 Db 5 riqcalaalslvialgvtgapsdprl---rqflqksalaa-----agkdelakyflae- 55  
 QY 64 PCLRWRPRKVGKPOLKAKE-----DLERSVDNLP---PRERRACKNFYKGF 108





QY 108 FTSC 111  
Db 118 ftsc 121

RESULT 9

AAY69789  
ID AAY69789 standard; Protein: 140 AA.

XX AC AAY69789;  
DT 20-APR-2000 (first entry)  
DE MWPsp-MWPmp20-(His)6-EGF-TEV-Somatostatin-28.  
KW Fusion protein; Bacillus; cell wall protein; promoter; cleavage site;  
KW TEV protease.  
XX OS Synthetic.

XX JP11341991-A.

XX 14-DEC-1999.

XX 30-MAR-1999; 95JP-0089488.

XX 31-MAR-1998; 98JP-0087339.

XX (ITOH-) ITOHAM FOODS INC.

XX (UDAK/) UDAKA S.

XX Sato S, Higashikuni N, Kudo T, Kondo M;  
WPI: 2000-101697/09.

XX N-PSDB; AA259209.

XX A DNA coding a new fused protein and preparation of a useful peptide  
through its expression.  
XX Example 6; Fig 3; 43pp; Japanese.

XX The invention relates to a DNA construct encoding a fusion protein  
comprising a bacillus species cell wall protein fused to a cleavage  
peptide and a heterologous protein. The fusion construct is placed  
downstream of a bacillus species promoter sequence. This sequence  
represents the fusion protein MWPsp-MWPmp20-(His)6-EGF-TEV-  
somatostatin-28, an example of the fusion protein of the invention.  
It comprises the bacillus brevis middle wall protein mp20 linked to  
the human somatostatin 28 protein via a (His)6-EGF linker-TEV protease  
cleavage site sequence.

XX Sequence 140 AA;

Query Match 17.7%; Score 105.5; DB 21; Length 140;  
Best Local Similarity 28.0%; Pred. No. 0.00012;  
Matches 42; Conservative 18; Mismatches 35; Indels 55; Gaps 8;

QY 4 SQHICALALIGLALATCSGAA--SQPDLDLASRLIQRALAAALPHRSGVSRWRTPYP 61

Db 4 svlasalaltvapmafaaeaaattapkmadmektvhh-----hhnsds----- 51

QY 62 MCP-----CL-----RWRPRKVKGPOLK-----AKEDLE 85

Db 52 -cplshdgcylhdgvcmylealdkycncvgyigercyrdlkweirdydipttenly 110

QY 86 -RVDNLP---PRERKACKNFYKGTSC 111

Db 111 fqsanspamaprerkagcknfwkftsc 140

RESULT 10

AAP20198

ID AAP20198 standard; Protein: 25 AA.

XX AC AAP20198;

XX 14-AUG-1992 (first entry)

XX Sequence of somatostatin-25 analogue.

XX Somatostatin; hormone; growth hormone release; inhibition.

XX Key Location/Qualifiers

FT Modified-site 1 /label= H-S

FT Disulfide-bond 14..25

FT Modified-site 25

FT /label= C-OH

XX US4316891-A.

XX 23-FEB-1982.

XX 14-JUN-1980; 80US-0159801.

XX 14-JUN-1980; 80US-0159801.

XX (SALK-) SALK INST BIOLOG.

XX Guillemin RCL, Esch FS, Bohlén P, Brazeau PE, Ling NC;  
WPI: 1982-19801E/10.

XX Extended somatostatin analogues-- with increased inhibition of  
growth hormone release

XX Claim 6; Column 12; 7pp; English.

XX The inventors claim a pharmaceutical compsn. which comprises (i)  
synthetic somatostatin-28 (SS-28), SS-25 or (D-Trp(22))-SS-28 and  
(ii) a liq. or solid carrier, and SS-28 deriva. and SS-25 deriva.  
The compsn. and deriva. are more potent than somatostatin in  
inhibiting release of growth hormone; they also inhibit basal and  
stimulated insulin and glucagon secretion. (D-Trp(22))-SS-28  
exhibits very substantial increases in potency w.r.t. inhibition of  
growth hormone secretion.

XX Sequence 25 AA;

Query Match 17.6%; Score 105; DB 3; Length 25;  
Best Local Similarity 81.0%; Pred. No. 1.7e-05;  
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 91 LPPRERKACKNFYKGTSC 111

Db 5 maprerkgcknfwkftsc 25

RESULT 11

AAB91017  
ID AAB91017 standard; Peptide: 25 AA.

XX AAB91017;

XX 22-JUN-2001 (first entry)

XX Somatostatin related peptide SEQ ID NO:191.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
blood component; modification; succinimidyl; maleimido group; amino;  
hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

OS Synthetic.

XX WO200069900-A2.  
PN 23-NOV-2000.  
XX 17-MAY-2000; 2000WO-US13576.  
XX 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0134406.  
PR 15-OCT-1999; 99US-0139783.  
XX (CONJ-) CONJUCHEM INC.  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
PI WPI; 2001-112059/12.  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
PT  
XX Disclosure; Page 252; 733pp; English.  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 25-AA;  
SQ

Query Match 17.6%; Score 105; DB 22; Length 25;  
Best Local Similarity 81.0%; Pred. No. 1.7e-05;  
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Y 91 LPPRRKAGCKNFYKGFSTC 111  
: ||||| ||||| |||||  
Db 5 maprerkagcknffwktftsc 25

RESULT 12  
AAW51859  
ID AAW51859 standard; peptide; 28 AA.  
XX  
XX AAW51859;  
XX 10-SEP-1998 (first entry)  
XX  
XX Somatostatin analogue for the treatment of syndrome X of Reaven.  
DE  
DE Somatostatin analogue; syndrome X of Reaven; hyperinsulinaemia syndrome;  
KW diazoxide; cyclothiazide; metformin.  
XX  
XX Synthetic.  
XX  
XX WO9810786-A2.  
PN  
XX 19-MAR-1998.  
PD  
XX 10-SEP-1997; 97WO-IL00301.  
PF

XX 10-OCT-1996; 96IL-0119403.  
PR 12-SEP-1996; 96IL-0119250.  
XX  
XX (COHE/) COHEN Y.  
XX Cohen Y;  
XX WPI; 1998-271636/24.  
DR  
XX Composition for treatment of the risk factors of syndrome X of  
PT Reaven - (hyperinsulinaemia syndrome) comprises somatostatin,  
PT diazoxide, cyclothiazide (or their analogues) and/or metformin  
XX  
XX Claim 42; Page 41; 45pp; English.  
XX The invention relates to a pharmaceutical composition for treatment of  
CC the risk factors of syndrome X of Reaven (hyperinsulinaemia syndrome). It  
CC comprises somatostatin, diazoxide, cyclothiazide (or an analogue of one  
CC of these) or metformin as the active ingredient. The composition reduces  
CC resistance to insulin, and so treats and prevents all the associated risk  
CC factors at once. The risk factors are hypertension, dyslipidaemia  
CC (raised triglyceride and LDL levels with reduced HDL levels), shorter  
CC coagulation time due to increased Plasminogen Activator Inhibitor-1  
CC levels, core obesity, glucose intolerance hyperinsulinaemia. The  
CC composition reduces the incidence of ischaemic heart disease, the  
CC cerebrovascular disorders, intermittent claudication, ischaemic bowel  
CC disease, impotence due to peripheral vascular disease, hypercoagulation  
CC (e.g. renal vein thrombosis), obesity and glucose intolerance. The  
CC present sequence represents a specifically claimed somatostatin analogue.  
XX  
XX Sequence 28 AA;  
SQ

Query Match 17.6%; Score 105; DB 19; Length 28;  
Best Local Similarity 81.0%; Pred. No. 2e-05;  
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Y 91 LPPRRKAGCKNFYKGFSTC 111  
: ||||| ||||| |||||  
Db 8 maprerkagcknffwktftsc 28

RESULT 13  
AAW28703  
ID AAW28703 standard; peptide; 28 AA.  
XX  
XX AAW28703;  
XX 07-OCT-1999 (first entry)  
XX  
XX Mouse somatostatin SS-28 hormone.  
XX  
XX Mouse somatostatin SS-28 hormone; growth hormone; insulin; glucagon;  
KW thyroid stimulating hormone; octreotide; cell-based delivery of insulin;  
KW glucose-stimulated insulin secretion; SSTAV; somatostatin receptor;  
KW mouse somatostatin receptor type V gene; diabetes.  
XX  
XX Mus musculus.  
OS  
XX WO9935242-A1.  
PN  
XX 15-JUL-1999.  
PD  
XX 11-JAN-1999; 99WO-US00633.  
PF  
XX 03-JUN-1998; 98US-0087848.  
PR 12-JAN-1998; 98US-0071193.  
PR 12-JAN-1998; 98US-0071209.  
PR 12-JAN-1998; 98US-0072556.  
PR 03-JUN-1998; 98US-0087821.  
XX  
XX (BETA-) BETAGENE INC.  
PA

XX Clark SA, Quaade C;  
XX WPI; 1999-444195/37.  
XX New defined medium for culture of neuroendocrine cells, e.g. of  
XX insulin-secreting cells  
XX  
XX Example 8; Page 143; 312pp; English.  
XX The present sequence is a mouse somatostatin (SS-28) hormone which  
XX was found to inhibit the release of growth hormone, thyroid stimulating  
XX hormone, insulin and glucagon. In addition, SS-28 and its analogue  
XX Octreotide may inhibit growth of some tumours. The hormone was  
XX used to study its effect on glucose-stimulated insulin secretion in  
XX high expressing and non-expressing clones of mouse somatostatin  
XX receptor, type V gene (SS28V). The insulin secretion was highly inhibited  
XX in the high expressing clone as compared to the non-expressing clone  
XX because the high expressing clone showed high sensitivity to  
XX somatostatin. The hormone effectively inhibits insulin secretion in the  
XX absence of glucose. The somatostatin receptor can be introduced in cell  
XX lines, used in cell-based delivery of insulin for treating diabetes, for  
XX precise regulation of insulin release.  
XX  
XX Sequence 28 AA;

Query Match 17.6%; Score 105; DB 20; Length 28;  
Best Local Similarity 81.0%; Pred. No. 2e-05; 2; Mismatches 0; Gaps 0;  
Matches 17; Conservative 2; Indels 0; Gaps 0;

QY 91 LPPRRKAGCKNFYWKGTSC 111  
DB 8 maprrkagcknfwtftsc 28

RESULT 14  
AY24384  
ID AAY24384 standard; peptide; 28 AA.

XX AC AAY24384;  
XX 20-SEP-1999 (first entry)  
XX Somatostatin peptide hormone SS-28.  
XX  
XX Glucagon-like peptide I receptor; GLP-1 receptor; drug screening;  
XX secretory function; immortalised neuroendocrine secretory cell;  
XX regulation; diabetes; insulin secretion; neuroendocrine-based disorder;  
XX Parkinson's disease; athyrotic cretinism; Addison's disease.  
XX  
XX Mus musculus.  
XX  
XX WO9935495-A2.  
XX  
XX 15-JUL-1999.  
XX  
XX 11-JAN-1999; 99WO-US00551.  
XX  
XX 03-JUN-1998; 98US-0087848.  
XX 12-JAN-1998; 98US-0071193.  
XX 12-JAN-1998; 98US-0071209.  
XX 12-JAN-1998; 98US-0072556.  
XX 03-JUN-1998; 98US-0087821.  
XX  
XX (BETA-) BETAGENE INC.  
XX  
XX Clark SA, Quaade C, Thigpen AE;  
XX WPI; 1999-430454/36.  
XX  
XX New modulators of secretory function, used to control peptide  
XX secretion from cells in vivo or in vitro, specifically for treating

PT diabetes  
XX  
XX Example 8; Page 153-154; 309pp; English.  
XX  
XX The present invention describes a method for identifying modulators (I)  
XX of secretory function by treating an immortalised cell, having a stable  
XX secretory function, with a test compound and detecting any change in  
XX secretion caused by the compound. (I) are used to control secretion of  
XX polypeptides from cells, in vivo or in vitro. Specifically they are used  
XX for treating or preventing diabetes by regulation of insulin secretion,  
XX but can also be used in cases of other neuroendocrine-based disorders,  
XX such as Parkinson's disease, athyrotic cretinism and Addison's disease.  
XX The method uses engineered, immortalised cells that are available in  
XX large amounts, with a stable and predictable phenotype. They allow  
XX screening to be performed in vivo. The present sequence represents a  
XX somatostatin peptide hormone SS-28 used in an example from the present  
XX invention.  
XX  
XX Sequence 28 AA;

Query Match 17.6%; Score 105; DB 20; Length 28;  
Best Local Similarity 81.0%; Pred. No. 2e-05; 2; Mismatches 0; Gaps 0;  
Matches 17; Conservative 2; Indels 0; Gaps 0;

QY 91 LPPRRKAGCKNFYWKGTSC 111  
DB 8 maprrkagcknfwtftsc 28

RESULT 15  
AY24239  
ID AAY24239 standard; peptide; 28 AA.

XX AC AAY24239;  
XX 15-SEP-1999 (first entry)  
XX Peptide hormone somatostatin SS-28.  
XX  
XX Glucagon-like peptide I receptor; GLP-1; regulation; secretion;  
XX neuroendocrine cell line; glycaemic sensing mechanism; glucose;  
XX genetic engineering; hypoglycaemia; diabetes.  
XX  
XX Synthetic.  
XX  
XX WO9935255-A2.  
XX  
XX 15-JUL-1999.  
XX  
XX 11-JAN-1999; 99WO-US00631.  
XX  
XX 03-JUN-1998; 98US-0087848.  
XX 12-JAN-1998; 98US-0071193.  
XX 12-JAN-1998; 98US-0071209.  
XX 12-JAN-1998; 98US-0072556.  
XX 03-JUN-1998; 98US-0087821.  
XX  
XX (BETA-) BETAGENE INC.  
XX  
XX Clark SA, Thigpen AE;  
XX WPI; 1999-419351/35.  
XX  
XX New immortalized neuroendocrine cells that stably secrete  
XX polypeptide, particularly used to treat diabetes and hypoglycaemia  
XX  
XX Example 8; Page 148; 318pp; English.  
XX  
XX The present invention describes immortalized neuroendocrine cells (A)  
XX that stably secrete a polypeptide hormone (I) contain an expression  
XX region that includes a transgene (TG), linked to a promoter functional  
XX in eukaryotic cells, such that expression of TG increases sensitivity

CC of the cells to a modulator of (I) secretion. (A) are specifically  
 CC used, by transplantation, to treat diabetes or hypoglycaemia (especially  
 CC where associated with insulin therapy) but more generally are used to  
 CC express, in vivo, a wide range of therapeutic hormones, enzymes,  
 CC amidated proteins and growth factors. Also engineered neuroendocrine  
 CC cells are used to identify new therapeutic agents or drug targets.  
 CC (A) have a stable phenotype and particularly inducible glucagon  
 CC secretion and glucose counter-regulatory capacities, i.e. they balance  
 CC the hyperglycaemic effects of beta-cell loss and the hypoglycaemic  
 CC effects of administered insulin. Since they are of human origin, they  
 CC are less likely to suffer immune rejection than xenografts. The present  
 CC sequence represents the peptide hormone somatostatin SS-28 used in an  
 CC example from the present invention.  
 XX  
 SQ Sequence 28 AA;

Query Match 17.6%; Score 105; DB 20; Length 28;  
 Best Local Similarity 81.0%; Pred. No. 2e-05;  
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 / 91 LPPRRKAGCKNFYWKGFSC 111  
 : ||||| ||||| |||||  
 Db 8 maprrkagcknffwktftsc 28

Search completed: June 13, 2002, 12:22:21  
 Job time: 274 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:55 ; Search time 28.68 Seconds  
(without alignments)  
94.534 Million cell updates/sec

Title: US-09-727-739B-15  
 Perfect score: 597  
 Sequence: 1 MRVSOIHCALALGLALAI...PPRRKAGCKNFWKGMSC 111

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628.seqs, 24425594 residues

total number of hits satisfying chosen parameters: 231628

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45

```

```
Database :
1: /cgn2_6/pdata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/pdata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/pdata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/pdata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/pdata/2/1aa/PCRU3_COMB.pep.*
6: /cgn2_6/pdata/2/1aa/backfiles1.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query No.	Length	DB	ID	Description
1	160.5	26.9	110	3	US-08-648-322-3	Sequence 3, Appl
2	87	14.6	112	3	US-08-648-322-2	Sequence 2, Appl
3	87	14.6	112	4	US-09-001-472-3	Sequence 3, Appl
4	83	13.9	85	3	US-08-648-322-6	Sequence 6, Appl
5	80.5	13.5	105	4	US-09-001-472-2	Sequence 2, Appl
6	79.5	13.3	109	3	US-08-648-322-5	Sequence 5, Appl
7	78	13.1	29	3	US-08-648-322-7	Sequence 7, Appl
8	78	13.1	29	3	US-08-648-322-11	Sequence 11, Appl
9	78	13.1	84	3	US-08-648-322-10	Sequence 10, Appl
10	77.5	13.0	943	1	US-08-455-370A-10	Sequence 10, Appl
11	77	12.9	14	1	US-07-977-638A-1	Sequence 1, Appl
12	77	12.9	14	1	US-08-235-272-5	Sequence 5, Appl
13	77	12.9	14	1	US-08-416-007-4	Sequence 4, Appl
14	77	12.9	14	1	US-08-676-263-11	Sequence 11, Appl
15	77	12.9	14	1	US-08-286-748B-13	Sequence 13, Appl
16	77	12.9	14	1	US-08-690-090A-1	Sequence 1, Appl
17	77	12.9	14	2	US-08-488-159-1	Sequence 1, Appl
18	77	12.9	14	2	US-08-465-764-1	Sequence 1, Appl
19	77	12.9	14	2	US-08-435-751-4	Sequence 4, Appl
20	77	12.9	14	2	US-08-282-360B-1	Sequence 1, Appl
21	77	12.9	14	2	US-08-747-137-13	Sequence 13, Appl
22	77	12.9	14	2	US-09-039-062-1	Sequence 1, Appl
23	77	12.9	14	2	US-09-042-224-1	Sequence 1, Appl
24	77	12.9	14	2	US-09-042-315A-1	Sequence 1, Appl
25	77	12.9	14	3	US-08-931-095-1	Sequence 1, Appl
26	77	12.9	14	4	US-09-100-414B-83	Sequence 83, Appl
27	77	12.9	14	4	US-09-420-866-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
 US-08-648-322-3  
 : Sequence 3, Application US/08648322  
 : Patent NO. 6074872  
 : GENERAL INFORMATION:  
 : APPLICANT: Sutcliffe, Gregor J.  
 : APPLICANT: de Lecea, Luis  
 : TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS  
 : NUMBER OF SEQUENCES: 24  
 : CORRESPONDENCE ADDRESS:  
 : 'ADDRESSEE: THE SCRIPES RESEARCH INSTITUTE  
 : STREET: 10666 No.. 6074872th Torrey Pines Road, TPC-8  
 : CITY: La Jolla  
 : STATE: California  
 : COUNTRY: US  
 : ZIP: 92037  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/648,322  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Fitting, Thomas  
 : REGISTRATION NUMBER: 34,163  
 : REFERENCE/DOCKET NUMBER: 519.0  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (619) 554-2937  
 : TELEFAX: (619) 554-6312  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 110 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : FRAGMENT TYPE: C-terminal  
 : US-08-648-322-3

	Query Match	26.9%	Score 160.5	DB 3	Length 110
	Best Local Similarity	38.3%	Pred. No. 8.e-13		
	Matches 46	Conservative 17	Mismatches 30	Indels 27	Gaps 6
QY	8	C A L A L G L A I A T C S - Q G A A S Q P D L D L R R L L Q A L A A A A L P H R S G S E R W R T Y P N C P C L	66		
DB	2	C A L A A C T I V I A L G V T Y G A P S D P R L - - - - R O F L A K S I L A A - - - - T G K E L A R Y F L A E - - - L	50		







```

1 GENERAL INFORMATION:
2 APPLICANT: Sutcliffe, Gregor J.
3 APPLICANT: de Lecea, Luis
4 TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
5 TITLE OF INVENTION: COMPOSITIONS AND METHODS
6 NUMBER OF SEQUENCES: 24
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
9 STREET: 10666 NO. 60748/27th Torrey Pines Road, TPC-8.
10 CITY: La Jolla
11 STATE: California
12 COUNTRY: US
13 ZIP: 92037
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/648,322
22 FILING DATE:
23
24 CLASSIFICATION: 435
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Fitting, Thomas
27 REGISTRATION NUMBER: 34,163
28 REFERENCE/DOCKET NUMBER: 519.0
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (619) 554-2937
31 TELEFAX: (619) 554-6312
32 INFORMATION FOR SEQ ID NO: 7:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 29 amino acids
35 TYPE: amino acid
36 TOPOLOGY: linear
37 MOLECULE TYPE: protein
38 FRAGMENT TYPE: C-terminal
39
40 US-08-648-322-7
41
42 Query Match 13.18; Score 78; DB 3; Length 29
43 Best Local Similarity .51.98; Pred. NO. 0.0035;
44 Matches 14; Conservative 2; Mismatches 11; Indels
45
46 QY 85 ERSVDNLPFRKAGCKNFYKGFSC 111
47 11 11 1111111111
48 DB 2 ERPPLOQPPHRDKKCKNFYKGFSSC 28
49
50 -SULT 8
51 -08-648-322-11
52 Sequence 11, Application US/08648322
53 Patent No. 6074872
54
55 GENERAL INFORMATION:
56 APPLICANT: Sutcliffe, Gregor J.
57 APPLICANT: de Lecea, Luis
58 TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
59 TITLE OF INVENTION: COMPOSITIONS AND METHODS
60 NUMBER OF SEQUENCES: 24
61 CORRESPONDENCE ADDRESS:
62 ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
63 STREET: 10666 NO. 60748/27th Torrey Pines Road, TPC-8.
64 CITY: La Jolla
65 STATE: California
66 COUNTRY: US
67 ZIP: 92037
68
69 COMPUTER READABLE FORM:
70 MEDIUM TYPE: Floppy disk
71 COMPUTER: IBM PC compatible
72 OPERATING SYSTEM: PC-DOS/MS-DOS
73 SOFTWARE: PatentIn Release #1.0, Version #1.25
74 CURRENT APPLICATION DATA:
75 APPLICATION NUMBER: US/08/648,322
76 FILING DATE:

```



APPLICATION NUMBER: US/08/255,272  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30742  
REFERENCE/DOCKET NUMBER: 6794-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-255-272-5

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.0018;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYKGTSC 111  
Db 1 AGCKNFFWKFTSC 14

RESULT 13  
US-08-416-007-4  
Sequence 4, Application US/08416007  
Patent No. 5693679  
GENERAL INFORMATION:  
APPLICANT: Vincent, Jean-Pierre  
APPLICANT: Gaudriault, Georges  
APPLICANT: Baudouin, Jean  
TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/08/416,007  
FILING DATE: 04-APR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06942/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-416-007-4

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.0018;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYKGTSC 111  
Db 1 AGCKNFFWKFTSC 14

RESULT 14  
US-08-676-263-11  
Sequence 11, Application US/08676263  
Patent No. 5705143  
GENERAL INFORMATION:  
APPLICANT: Bower, Gary R.  
APPLICANT: Forster, Alan M.  
APPLICANT: Riley, Anthony L. M.  
APPLICANT: Storey, Anthony E.  
TITLE OF INVENTION: BIOLOGICAL TARGETING AGENTS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/08/676,263  
FILING DATE: 07-NOV-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94300224.6  
FILING DATE: 12-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: (312) 474-6600  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Cross-links  
LOCATION: 3:14  
US-08-676-263-11

Query Match 12.9%; Score 77; DB 1; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.0018;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYKGTSC 111  
Db 1 AGCKNFFWKFTSC 14

RESULT 15  
US-08-286-748B-13  
Sequence 13, Application US/08286748B  
Patent No. 5759542

```

GENERAL INFORMATION:
APPLICANT: Victor Gurewich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
OF DRUGS BY PLATELETS FOR THE TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-13

```

```

Query Match 12.9%; Score 77; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.0018;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Caps 0;
OY 98 AGCKNFYWKGTSC 111
1 AGCKNFYWKGTSC 14

```

Search completed: June 13, 2002, 12:22:56  
Job time: 194 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using .sw model

Run on: June 13, 2002, 12:29:53 ; Search time 18.73 seconds  
(without alignments) 229.464 Million cell updates/sec

Title: US-09-727-739B-15  
 Perfect score: 597  
 Sequence: 1 MRVSIHCALALLGLALAIC.....PPRRKAGCKNFYWKGTSC 111

```

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5

Searched:          105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

```

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	1	476	79.7	115	1	SMS2_ONCMY	Q31194 oncorhynch
2	2	205.5	34.4	120	1	SMS2_CARAU	Q3Y9h4 carassius a
3	3	189	31.7	135	1	SMS2_LOPHU	P01170 lophius ame
4	4	179.5	30.1	116	1	SMS2_CHICK	P33094 gallus gall
5	5	178.5	29.9	114	1	SMSA_CARAU	Q3Y9h5 carassius a
6	6	176.5	29.6	116	1	SMS_HUMAN	P01166 homo sapien
7	7	174	29.1	115	1	SMS1_RANRI	P87384 rana ridibu
8	8	166.5	27.9	116	1	SMS_SHEEP	O46688 ovis aries
9	9	163.5	27.4	116	1	SMS_MOUSE	P01167 mus musculu
10	10	162.5	27.2	116	1	SMS_BOVIN	P26917 bos taurus
11	11	160.5	26.9	111	1	SMS_ICTPU	P49670 canis fami
12	12	156.5	26.2	114	1	SMS1_CARAU	P01171 ictalurus p
13	13	153	25.6	115	1	SMS1_PROAN	Q3W7f0 protopterus
14	14	136	22.8	73	1	SMS2_PLAFE	P21780 platichthys
15	15	136	22.8	92	1	SMS_PIG	Q3W7f0 protopterus
16	16	131.5	22.0	121	1	SMS1_LOPAM	P01168 sus scrofa
17	17	125.5	21.0	74	1	SMS2_MYOSC	P01169 lophius ame
18	18	121	20.3	28	1	SMS2_ORONI	Q3W7f0 protopterus
19	19	107	17.9	26	1	SMS1_AMICA	P01168 sus scrofa
20	20	104	17.4	34	1	SMS1_MYXGL	P01169 lophius ame
21	21	102.5	17.2	111	1	SMSB_CARAU	Q3P829 oreochromis
22	22	106.5	16.8	109	1	SMS2_PETNA	Q3P829 oreochromis
23	23	94	15.7	37	1	SMS2_PETNA	Q3P829 oreochromis
24	24	93	15.6	35	1	SMS1_LAMFL	P19205 myxine glut
25	25	91.5	15.3	103	1	SMS2_RANRI	P19205 myxine glut
26	26	87	14.6	112	1	CORT_RAT	Q3Y9h3 carassius a
27	27	80.5	13.5	105	1	CORT_HUMAN	Q3Y9h3 carassius a
28	28	79.5	13.3	109	1	CORT_MOUSE	Q3W7f0 protopterus
29	29	77.5	13.0	1236	1	POLS_WEEV	P21779 petromyzon
30	30	77	12.9	14	1	SMS1_MYOSC	Q3P829 oreochromis
31	31	77	12.9	14	1	SMS1_MYOSC	Q3P829 oreochromis
32	32	74.5	12.5	234	1	SMS_ALAMI	Q3Y9h3 carassius a
33	33	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
34	34	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
35	35	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
36	36	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
37	37	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
38	38	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
39	39	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
40	40	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
41	41	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
42	42	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
43	43	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
44	44	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
45	45	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
46	46	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
47	47	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
48	48	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
49	49	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
50	50	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
51	51	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
52	52	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
53	53	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
54	54	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
55	55	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
56	56	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
57	57	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
58	58	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
59	59	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
60	60	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
61	61	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
62	62	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
63	63	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
64	64	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
65	65	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
66	66	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
67	67	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
68	68	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
69	69	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
70	70	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
71	71	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
72	72	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
73	73	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
74	74	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
75	75	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
76	76	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
77	77	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
78	78	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
79	79	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
80	80	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
81	81	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
82	82	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
83	83	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
84	84	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
85	85	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
86	86	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
87	87	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
88	88	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
89	89	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
90	90	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
91	91	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
92	92	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
93	93	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
94	94	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
95	95	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
96	96	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
97	97	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
98	98	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
99	99	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
100	100	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
101	101	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
102	102	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
103	103	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
104	104	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
105	105	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
106	106	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
107	107	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
108	108	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
109	109	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
110	110	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
111	111	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
112	112	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
113	113	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
114	114	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
115	115	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
116	116	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
117	117	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
118	118	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
119	119	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
120	120	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
121	121	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
122	122	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
123	123	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
124	124	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
125	125	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
126	126	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
127	127	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
128	128	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
129	129	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
130	130	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
131	131	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
132	132	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
133	133	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
134	134	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
135	135	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
136	136	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
137	137	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
138	138	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
139	139	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
140	140	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
141	141	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
142	142	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
143	143	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
144	144	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
145	145	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
146	146	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
147	147	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
148	148	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
149	149	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
150	150	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
151	151	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
152	152	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
153	153	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
154	154	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
155	155	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
156	156	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
157	157	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
158	158	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
159	159	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
160	160	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
161	161	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
162	162	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
163	163	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
164	164	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
165	165	73.5	12.3	342	1	CYS2_HABCO	P21779 petromyzon
166	166	73.5	12.3	342	1</		

34	70	11.7	1	Y4RB RHES	P55635	rhizomys
35	69.5	11.6	342	1	P19092	hemionchus
36	69.5	11.6	1305	1	P79784	rattus norv
37	69	11.6	1204	1	P29473	bos taurus
38	68.5	11.5	399	1	P34821	mus musculus
39	67	11.2	133	1	O09006	mus musculus
40	67	11.2	260	1	O13374	saccharomyc
41	66	11.1	1204	1	NS03_P3G	mus musculus
42	66	11.1	1204	1	NS03_P3G	mus musculus
43	64.5	10.8	317	1	CAH6 MOUSE	mus scrofa
44	64	10.7	105	1	SMS2 ICTPU	mus musculus
45	64	10.7	988	1	E4L2 MOUSE	mus musculus
46	62.5	10.5	100	1	APL2 MACFA	macaca fasci
47	62.5	10.5	100	1	APL2 MACFA	macaca fasci

## ALIGNMENTS

```

RESULT 1
SM52_ONCMY SM52_ONCMY STANDARD; PRT; 115 AA.
AC Q91194;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: {Tyr21,Gly24}somatostatatin-28;
DE [Tyr7,Gly10]somatostatatin-14].
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OC NBSI_Taxid=8022;
OX [.]
SRN SEQUENCE FROM N.A.
RRX MEDLINE=95354921; PubMed=7628684;
RRR Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;
RRR "Isolation and characterization of a cDNA encoding for
RRR preprosomatostatin containing [Tyr7, Gly10]somatostatatin-14 from the
RRR endocrine pancreas of rainbow trout, Oncorhynchus mykiss.";
RRR Gen. Comp. Endocrinol. 98:253-261(1995).
CCC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CCC -1- SUBCELLULAR LOCATION: Secreted.
CCC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
-----
CCC This SWISS-PROT entry is copyright. It is produced through a collaboration
CCC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CCC the European Bioinformatics Institute. There are no restrictions on its
CCC use by non-profit institutions as long as its content is in no way
CCC modified and this statement is not removed. Usage by and for commercial
CCC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CCC or send an email to license@isb-sib.ch).

```

```

Query Match      79.7%; Score 476; DB 1; Length 115;
Best Local Similarity 83.1%; Pred. No. 2e-44;
Matches 98; Conservative 3; Mismatches 7; Indels 10; Gaps 4;
2Y      1 MRYSTHCAALLGIALGIAICSGAASQDPLDLRRLQRLAALPHRSGVSEWTFY 60
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DDB      1 MRYCHCAALLGIALGIAICSGAASQDPLDLRRLQRLAALPHRSGVSEWTFY 60
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2Y      61 PNCPLRLWRPKVKGSPOLKAE---LERSV---DNLPRRKACGNFYWKGTSC 111
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 61 PNCPL--RPRVKCP-AGAKEDLVELERSVGNPNLPPRERKAGCNFTWKGTSC 115

RESULT 2  
SMS2\_CARAU STANDARD; PRT; 120 AA.

ID SMS2\_CARAU  
AC Q9IGH4; Q9PTU2;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr21, Gly24]somatostatin-28;  
DE [Tyr7, Gly10]somatostatin-14]  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Osteichthyes;  
OC Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Brain;  
RL Lin X-W. Peter R.E.;  
RL "Cloning and characterization of cDNAs encoding preprosomatostatin-I  
and -II from goldfish brain.";  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Liver;  
RL Otto C.J., Peter R.E.;  
RL "The expression of SRIF mRNA in the brain of goldfish.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch)  
CC  
CC EMBL: U60262; AAD09626.1;  
CC EMBL: AF025686; AAF15306.1;  
CC InterPro: IPR004250; Somatostatin.  
CC Pfam: PF03002; Somatostatin; 1.  
CC Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
CC SIGNAL 1 23 POTENTIAL.  
CC PROPEP 24 92 POTENTIAL.  
CC PEPTIDE 93 120 [TYR21, GLY24]SOMATOSTATIN-28.  
CC FT PEPTIDE 107 120 [TYR7, GLY10]SOMATOSTATIN-14.  
CC FT DISULFID 109 120 BY SIMILARITY.  
CC FT CONFLICT 51 51 Q -> RW (IN REF. 2).  
CC FT CONFLICT 51 51  
CC SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;  
CC  
CC Query Match 34.4%; Score 205.5; DB 1; Length 120;  
CC Best Local Similarity 42.0%; Pred. No. 2.6e-15;  
CC Matches 55; Conservative 14; Mismatches 31; Indels 31; Gaps 7;  
CC  
CC QY 1 MRYSTHCHALALLGLALALCSQGAASQ--PDLDLASRLQLQALAAALPHRSGVSRWT 58  
CC Db 1 MRCELCVYALLGLSLVLCRCANSLQEPDLDFRHLRLQRA-----SNTQATQD 52  
CC  
CC QY 59 FYP-----NCPCLRWPRVKGFQKAK-EDL-----ERSVDN---LPPPRKAGC 100  
CC Db 53 FTRDVEKLLSLSLIPMEMRE---KGLSMAGESDLRLQEQSAESSNQLPTRVKEGC 109  
CC  
CC QY 101 KNEYWKGFTSC 111  
CC Db 110 KNEYWKGFTSC 120

RESULT 3  
SMS2\_LOPAM STANDARD; PRT; 125 AA.

ID SMS2\_LOPAM  
AC P01170; Q91066;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr7, Gly10]somatostatin-14].  
DE Lophius americanus (American goosefish) (Anglerfish).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Osteichthyes;  
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.  
OX NCBI\_TaxID=8073;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81052423; PubMed=6107860;  
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;  
RA "Cloning and sequence analysis of cDNAs encoding two distinct  
somatostatin precursors found in the endocrine pancreas of  
anglerfish.";  
RT anglerfish.  
RN [2]  
RP PARTIAL SEQUENCE, AND HYDROXYLATION.  
RX MEDLINE=87308304; PubMed=2887572;  
RA Andrews P.C., Nichols R., Dixon J.E.;  
RA "Post-translational processing of preprosomatostatin-II examined  
using fast atom bombardment mass spectrometry.";  
RL J. Biol. Chem. 262:12692-12699(1987).  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF  
CC ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch)  
CC  
CC EMBL: V00641; CAA23987.1;  
CC FIR; A01434; KIAFS2.  
CC FIR; A27376; A27376.  
CC InterPro: IPR004250; Somatostatin.  
CC Pfam: PF03002; Somatostatin; 1.  
CC Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;  
CC Multigene family.  
CC SIGNAL 1 24 POTENTIAL.  
CC PROPEP 25 109 [TYR7, GLY10]SOMATOSTATIN-14.  
CC FT PEPTIDE 112 125 HYDROXYLATION.  
CC FT DISULFID 114 125 DV -> TG (IN REF. 1).  
CC FT MOD\_RES 120 120 G -> E (IN REF. 1).  
CC FT CONFLICT 77 78  
CC FT CONFLICT 90 90  
CC SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;  
CC  
CC Query Match 31.7%; Score 189; DB 1; Length 125;  
CC Best Local Similarity 39.7%; Pred. No. 1.6e-13;  
CC Matches 52; Conservative 13; Mismatches 32; Indels 34; Gaps 7;  
CC  
CC QY 6 IHCALALGLALALCSQGAASQ-----PDLDLASRR--LQALAAALPHRSGVSR 55  
CC Db 4 IRCPALALLALVLCGSPVSSQLDREQSDNODLRLHRLHRLRARSAGL-----LSQE 58  
CC  
CC QY 56 W-----RTFYPNCPCLRWPRVKGFQKAK--AKEDLERSVD---LPPPRKAGC 100  
CC Db 59 WSKRAVEELLQMSLPADV---QREADASMATGGRNRLERSVDSTNLPPRKAGC 114  
CC  
CC QY 101 KNEYWKGFTSC 111



DB 115 KNFYWKGFSC 125

RESULT 4  
SMS\_CHK ID SMS\_CHK STANDARD; PRT: 116 AA.

AC P33094;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
GN SST.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RA Yonekura H., Okamoto H., Kato S., Yamamoto H.,  
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X60191; CAA42747.1;  
CC PIR: S20630;  
CC InterPro: IPR004250; Somatostatin.  
CC Pfam: PF03002; Somatostatin; 1.  
CC Cleavage on pair of basic residues; Hormone; Signal.  
CC SIGNAL 1 24 BY SIMILARITY.  
CC PROPEP 25 88 BY SIMILARITY.  
CC PEPTIDE 89 116 SOMATOSTATIN-28.  
CC PEPTIDE 103 116 SOMATOSTATIN-14.  
CC DISULFID 105 116  
CC SEQUENCE 116 AA; 12675 MW; 8A5BB9BDA8A291BA CRC64;  
CC  
CC Query Match 30.1%; Score 179.5; DB 1; Length 116;  
CC Best Local Similarity 39.8%; Pred. No. 1.6e-12;  
CC Matches 49; Conservative 20; Mismatches 27; Indels 27; Gaps 6;  
CC  
DB 5 QIICALALLGLALALICSGA--SOPDLASRLRLQALAAALPHSGVSEWRFTYPC 63  
DB 5 RLQCALALLSLALVGTGVSAPSPRL---RQFLQSLAA--AGQELAKFLAE- 55  
DB 64 PCLWRPRKVGKPOLAKE-----DLERSVDNLP---PRERKAGCKNYWKG 108  
DB 56 --LLSEFSQTEALESEDSRGASQDEVRLELERSANSNPALPRERKAGCKNFWTF 113  
DB 109 TSC 111  
DB 114 TSC 116

RESULT 5  
SMS\_CARAU ID SMS\_CARAU STANDARD; PRT: 114 AA.  
AC Q9YGH5;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin IA precursor [Contains: Somatostatin-26; Somatostatin-14].

DE 14].  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Lin X.-W., Peter R.E.;  
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-1  
RT and -II from goldfish brain";  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U40754; AAD09359.1;  
CC InterPro: IPR004250; Somatostatin.  
CC Pfam: PF03002; Somatostatin; 1.  
CC Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
CC SIGNAL 1 24 POTENTIAL.  
CC PROPEP 25 88  
CC PEPTIDE 89 114 SOMATOSTATIN-26 (POTENTIAL).  
CC PEPTIDE 101 114 SOMATOSTATIN-14.  
CC DISULFID 103 114 BY SIMILARITY.  
CC SEQUENCE 114 AA; 12574 MW; B5920015E2D272A4 CRC64;  
CC  
CC Query Match 29.9%; Score 178.5; DB 1; Length 114;  
CC Best Local Similarity 38.2%; Pred. No. 2e-12;  
CC Matches 50; Conservative 15; Mismatches 29; Indels 37; Gaps 5;  
CC  
OY 1 MRVSOIHCALALIGLALALICSGAASOPDLASRLRLQALAAALPHSGVSEWRFTY 60  
DB 1 MLSTRIOCALALLSLALAVCSVSA--PTDAKLRLQLQSL--NPAKQE----- 47  
OY 61- PNCPLWRPRKVGKPOLKAK-----EDLERSVDN-----LPPRRKAGC 100  
DB 48 ----LARYTLADLLSELVQAEALPEDLGRAVEXDEVRLELRAAGPMLAPRRKAGC 103  
OY 101 KNFYWKGFSC 111  
DB 104 KNFYWKGFSC 114  
DB  
RESULT 6  
SMS\_HUMAN ID SMS\_HUMAN STANDARD; PRT: 116 AA.  
AC P01166; 1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
GN SST.  
OS Homo sapiens (Human), and  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606, 9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=Human;  
RX MEDLINE=84146798; PubMed=6142531;  
RA Shen L.-P., Rutter W.J.;



DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
 GN SPT.  
 OS Ovis aries (Sheep).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ILE DE FRANCE;  
 RX MEDLINE-99094691; PubMed-9880082;  
 RA Bruneau G., Tillet Y.;  
 RT "Localization of the preprosomatostatin-mRNA by in situ hybridization  
 in the ewe hypothalamus.";  
 RL Peptides 19:1749-1758(1998).  
 CC -|- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
 CC -|- SUBCELLULAR LOCATION: Secreted.  
 CC -|- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
 C  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF031488; AAC04697.1;  
 DR EMBL; Y15267; CA75556.1;  
 DR InterPro; IPR004250; Somatostatin.  
 DR Pfam; PF03002; Somatostatin; 1.  
 KW Cleavage on pair of basic residues; Hormone; Signal.  
 FT SIGNAL 1 24 BY SIMILARITY.  
 FT PROPEP 25 88 BY SIMILARITY.  
 FT PEPTIDE 89 116 SOMATOSTATIN-28.  
 FT PEPTIDE 103 116 SOMATOSTATIN-14.  
 FT DISULFID 105 116 BY SIMILARITY.  
 SQ SEQUENCE 116 AA; 12689 MW; C18F17E31A3718DE CRC64;  
  
 Query Match 27.9%; Score 166.5; DB 1; Length 116;  
 Best Local Similarity 37.4%; Pred. No. 4e-11;  
 Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;  
  
 QY 5 OIHCNALIGLALACS-QGRASQPDLDASRLQLQALAAALPHRSGVSEWRFTYFNC 63  
 Yb 5 RLQALAAALSVLALGGVTVGAPSDPRL-----RQFLQKSLAA-----AGKQELAKYFLAE- 55  
 YQ 64 PCLWRPRKVKGPQLKAKE-----DLERSVDNLP-----PRRKAGCKNFKYKGF 108  
 Db 56 --LLSEPNQTDNALEPEDLSQAQEDMRLEQLRSANSNPANAPRRKAGCKNFKWTF 113  
 QY 109 TSC 111  
 Yb 111  
 Db 114 TSC 116  
  
 RESULT 9  
 SMS\_MOUSE  
 ID SMS\_MOUSE STANDARD; PRT; 116 AA.  
 AC P01167;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Somatostatin precursor [Contains: Antrin; Somatostatin-28;  
 DE Somatostatin-14].  
 GN SST OR SMT.  
 OS Mus musculus (Mouse), and  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090, 10116;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-85006903; PubMed-6148343;  
 RA Tavianini M.A., Hayes T.E., Magazin M.D., Minth C.D., Dixon J.E.;  
 RT "Isolation, characterization, and DNA sequence of the rat  
 somatostatin gene.";  
 RL J. Biol. Chem. 259:11798-11803(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-83238516; PubMed-6134734;  
 RA Argos P., Taylor W.L., Minth C.D., Dixon J.E.;  
 RT "Nucleotide and amino acid sequence comparisons of  
 preprosomatostatins.";  
 RL J. Biol. Chem. 258:8788-8793(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-83213516; PubMed-6133871;  
 RA Goodman R.H., Aron D.C., Roos B.A.;  
 RT "Rat pre-prosomatostatin. Structure and processing by microsomal  
 membranes.";  
 RL J. Biol. Chem. 258:5570-5573(1983).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-85303984; PubMed-28693939;  
 RA Goodman R.H., Montminy M.R., Low M.J., Habener J.F.;  
 RT "Biosynthesis of rat preprosomatostatin.";  
 RL Adv. Exp. Med. Biol. 188:31-47(1985).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-84221954; PubMed-6145156;  
 RA Montminy M.R., Goodman R.H., Horovitch S.J., Habener J.F.;  
 RT "Primary structure of the gene encoding rat preprosomatostatin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3337-3340(1984).  
 RN [6]  
 RP SEQUENCE OF 38-116 FROM N.A.  
 RC SPECIES-Rat;  
 RX MEDLINE-82120034; PubMed-6120163;  
 RA Goodman R.H., Jacobs J.W., Dee P.C., Habener J.F.;  
 RT "Somatostatin-28 encoded in a cloned cDNA obtained from a rat  
 medullary thyroid carcinoma.";  
 RL J. Biol. Chem. 257:1156-1159(1982).  
 RN [7]  
 RP SEQUENCE OF 25-34.  
 RC SPECIES-Rat; STRAIN-SPRAGUE-DAWLEY;  
 RX MEDLINE-88070564; PubMed-2891188;  
 RA Benoit R., Ling N., Esch F.;  
 RT "A new prosomatostatin-derived peptide reveals a pattern for  
 prohormone cleavage at monobasic sites.";  
 RL Science 238:1126-1129(1987).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-Mouse; TISSUE-Brain;  
 RX MEDLINE-90206793; PubMed-1969620;  
 RA Fuhrmann G., Heilig R., Kempf J., Ebel A.;  
 RT "Nucleotide sequence of the mouse preprosomatostatin gene.";  
 RL Nucleic Acids Res. 18:1287-1287(1990).  
 CC -|- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
 CC -|- SUBCELLULAR LOCATION: Secreted.  
 CC -|- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
 C  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

```

DR EMBL; K02248; AAA42161.1; -
DR EMBL; V01271; CAA24579.1; -
DR EMBL; J00787; AAA42164.1; -
DR EMBL; M25890; AAA42167.1; -
DR EMBL; J00788; AAA42162.1; -
DR EMBL; X51458; CAA35831.1; -
DR PIR; A20983; RIRIS1; -
DR PIR; S08416; S08416; -
DR MGI; 98326; Smst.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal.
FT SIGNAL 1 24 ANTRIN.
FT PROPEP 25 34
FT PEPTIDE 35 88
FT PEPTIDE 89 116
FT PEPTIDE 103 116
FT DISULFID 105 116
FT CONFLICT 43 43
FT CONFLICT 79 79
FT SEQUENCE 116 AA; 12745 MW; D48B5454C4490375 CRC64;

Query Match 27.4%; Score 163.5; DB 1; Length 116;
Best Local Similarity 37.4%; Pred. No. 8.3e-11;
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;

QY 5 QTHCALALIGLALAICS-OGNASOPDLASRLRLQALAAALPHRSVGRWTFYPNC 63
DB 5 RIQCALALSVILAGVGTGAPDRL-----RQFLQKSLAAA-----TKQELAKYFLAE- 55
QY 64 PCLWRPRKVGPKLKAKE-----DLERSVDMLP---PREKACKNFYKGF 108
DB 56 --LLSEPNOTENDALEPEDLPQAAQDEMRLQRSANSNPAMAPRRKACKNFFWKTF 113
QY 109 TSC 111
DB 114 TSC 116

RESULT 10
SMS_BOVIN
ID SMS_BOVIN STANDARD; PRT; 116 AA.
AC P26917;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
GN SST
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88288237; PubMed=2899837;
RA Su C.J., White J.W., Li W.H., Luo C.C., Frazier M.L., Saunders G.F.,
RA Chan L.;
RT "Structure and evolution of somatostatin genes.";
RL Mol. Endocrinol. 2:209-216(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=99198780; PubMed=10100681;
RA Furu L.M., Kazmer G.W., Strausbaugh L., Zinn S.A.;
RT "Cloning and characterization of the bovine somatostatin gene.";
RL J. Anim. Sci. 77:492-493(1993).
CC 1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31217; AAA30744.1; -
DR EMBL; U97077; AAB58056.1; -
DR PIR; A40929; A40929.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT PROPEP 25 88
FT PEPTIDE 89 116
FT PEPTIDE 103 116
FT DISULFID 105 116
FT SEQUENCE 116 AA; 12688 MW; C18F17E64A371D8E CRC64;

Query Match 27.2%; Score 162.5; DB 1; Length 116;
Best Local Similarity 37.4%; Pred. No. 1.1e-10;
Matches 46; Conservative 19; Mismatches 31; Indels 27; Gaps 6;

QY 5 QTHCALALIGLALAICS-OGNASOPDLASRLRLQALAAALPHRSVGRWTFYPNC 63
DB 5 RIQCALALSVILAGVGTGAPDRL-----RQFLQKSLAAA-----AGKQELAKYFLAE- 55
QY 64 PCLWRPRKVGPKLKAKE-----DLERSVDMLP---PREKACKNFYKGF 108
DB 56 --LLSEPNOTENDALEPEDLPQAAQDEMRLQRSANSNPAMAPRRKACKNFFWKTF 113
QY 109 TSC 111
DB 114 TSC 116

RESULT 11
SMS_CANFA
ID SMS_CANFA STANDARD; PRT; 116 AA.
AC P49670;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
GN SST
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97142297; PubMed=8988514;
RA Dickinson C.J., Delvalle J., Todisco A., Gantz I., Tong L.,
RA Finniss S., Yamada T.;
RT "Canine prosomatostatin: Isolation of a cDNA, regulation of gene
RT expression; and characterization of post-translational processing
RT intermediates.";
RL Regul. Pept. 67:145-152(1996).
CC 1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; L42325; AAA67099.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
DR SIGNAL 1 24 BY SIMILARITY.
FT SIGNAL 25 88 BY SIMILARITY.
FT PEPTIDE 89 116 SOMATOSTATIN-28.
FT PEPTIDE 103 116 SOMATOSTATIN-14.
FT DISULFID 105 116 BY SIMILARITY.
SQ SEQUENCE 116 AA; 12735 MW; AB49BD231E731C9E CRC64;

Query Match 26.9%; Score 160.5; DB 1; Length 116;
Best Local Similarity 36.5%; Pred. No. 1.8e-10;
Matches 46; Conservative 19; Mismatches 28; Indels 33; Gaps 7;

QY 5 OIHCALALGLALAI-----CSGGAASQPDLDASRLQLPALAAALPHRSGVSEWRTFY 60
Db 5 RLQALALSLVIALGGVTC---AFSDPRL---RGLQKSLAA-----AGQELAKYFL 53
Y 61 PNCPCPLRWPRKVKGPOLKAKE-----DLERSVDNLP---PRERKAGCKNFPY 105
Db 54 AE---LLSPFNQTDALPEDELQAAEQDEMLELQRSANSNPAMPAPRERKAGCKNFF 110
QY 106 KGFTSC 111
Db 111 KFTTSC 116

RESULT 12
SMS1 ICTPU STANDARD; PRT; 114 AA.
AC P01171;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin I precursor (Contains: Somatostatin-14 (SS-14)).
OS Ictalarus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Ictaluridae; Ictalurus.
OX NCBI_TaxId=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85303576; PubMed=2863931;
RA Dixon J.E., Andrews P.C.;
RA "Somatostatins of the channel catfish."
RA Adv. Exp. Med. Biol. 188:19-29(1985).
L N [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=82265698; PubMed=6179939;
RA Minth C.D., Taylor W.L., Magazini M.D., Taviani M.A., Collier K.J.,
RA Weith H.L., Dixon J.E.;
RA "The structure of cloned DNA complementary to catfish pancreatic
RA somatostatin-14 messenger RNA."
RA J. Biol. Chem. 257:10372-10377(1982).
RN [3]
RP SEQUENCE OF 82-114 FROM N.A.
RA MEDLINE=82082515; PubMed=6171821;
RA Taylor W.L., Collier K.J., Deschenes R.J., Weith H.L., Dixon J.E.;
RA "Sequence analysis of a cDNA coding for a pancreatic precursor to
RA somatostatin."
RA Proc. Natl. Acad. Sci. U.S.A. 78:6694-6698(1981).
RN [4]
RP SEQUENCE OF 101-114.
RA MEDLINE=81264223; PubMed=6114953;
RA Andrews P.C., Dixon J.E.;
RA "Isolation and structure of a peptide hormone predicted from a mRNA
RA sequence. A second somatostatin from the catfish pancreas."
RA J. Biol. Chem. 256:8267-8270(1981).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M25903; AAA49339.1;
DR EMBL; V00607; CAA23877.1;
DR EMBL; V00608; CAA23878.1;
DR PIR; S00292; RIIDS1
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
DR Cleavage on pair of basic residues; Hormone; Signal;
KW Multigene family. 24 PROBABLE
FT SIGNAL 101 114 SOMATOSTATIN-14.
FT PEPTIDE 103 114
FT DISULFID 62 62
FT CONFLICT 103 114 E -> Q (IN REF. 2).
SQ SEQUENCE 114 AA; 12419 MW; FEE0F2C76F74D99F CRC64;

Query Match 26.2%; Score 156.5; DB 1; Length 114;
Best Local Similarity 31.4%; Pred. No. 4.6e-10;
Matches 43; Conservative 20; Mismatches 25; Indels 49; Gaps 5;

QY 1 MRYVQTHCALALLGLALAICS--QGAASQPDLDASRLQLPALAAALPHRSGVSEWRTF 59
Db 1 MPSTRIQCALALLAVALSVCSVGAPS-----DAKLQFLQRSILA----- 41
QY 60 YPNCPCPLRWPRKVKGPOLKAK-----EDLERSVDN-----LPPR 94
Db 42 ----PSVQQLRYTLAELLAELAENELVDSEVSRRAESGARLEMERAGPMLAPR 97
QY 95 ERKAGCKNFPYKGFYSC 111
Db 98 ERKAGCKNFPYKGFYSC 114

RESULT 13
SMS1 PROAN STANDARD; PRT; 115 AA.
AC Q9W7F0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin I precursor (PSS1) (Contains: Somatostatin-27;
DE Somatostatin-14)
OS Protopterus annectens (African lungfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Dipnoi; Lepidosireniformes; Protopterae; Protopterus.
OX NCBI_TaxId=7888;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA MEDLINE=99326690; PubMed=10398054;
RA Trabucchi M., Tostivint H., Lihmann I., Jegou S., Vallarino M.,
RA Vaudry H.;
RA "Molecular cloning of the cDNAs and distribution of the mRNAs encoding
RA two somatostatin precursors in the African lungfish Protopterus
RA annectens."
RA J. Comp. Neurol. 410:643-652(1999).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```

DR PIR: A01432; RIGGS.
DR PIR: A24222; A24222.
DR PIR: A34109; A34109.
DR PIR: S13616; S13616.
DR InterPro: IPR004250; Somatostatin.
DR Pfam: PF03002; Somatostatin_1.
KW Cleavage on pair of basic residues; Hormone.
FT NONTER 1
FT PROPEP 1
FT PEPTIDE 65
FT PEPTIDE 79
FT DISULFID 81
SQ SEQUENCE 92 AA; 10346 MW; 787CB82CFBBAE76 CRC64;

Query Match 22.8%; Score 136; DB 1; Length 92;
Best Local Similarity 36.9%; Pred. No. 5.9e-08;
Matches 38; Conservative 15; Mismatches 24; Indels 26; Gaps 5;

y 24 AASQDLDLASRLQALAAALPHRSGVSEWRWTFYPCPCLRWRPRKVGPKAKE- 82
db 1 APSDRL-----RQFLQKSLAAA-----AGQELAKYFLAE---LLSEPNQTENDALEPEDL 49
Oy 83 -----DLERSVDNLP---PRERKAGCKNFYKGTSC 111
Db 50 SQAQEDQEMRLQRSANSNPAMAPRRKAGCKNFYKGTSC 92

```

Search completed: June 13, 2002, 12:29:54  
 Job time: 446 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:31:02 ; Search time 60.72 Seconds  
(without alignments)

Title: US-09-727-739B-18  
Perfect score: 58  
Sequence: 1 SVDNLPPRERK 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

**Database :**

1: sp:archaea:\*

2: sp:bacteria:\*

3: sp:fungi:\*

4: sp:human:\*

5: sp:inverteb:\*

6: sp:mammal:\*

7: sp:mhc:\*

8: sp:organell:\*

9: sp:phage:\*

10: sp:plant:\*

11: sp:rodent:\*

12: sp:virus:\*

13: sp:verteb:\*

14: sp:unclass:\*

15: sp:rvirus:\*

16: sp:bacteri:\*

17: sp:arcea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	46	79.3	115	13	Q9Y43	Q9Y43	osteoGLOSSu
2	46	79.3	845	5	Q9YA66	Q9YA66	droSOPhila
3	42	72.4	273	3	O13787	O13787	schizosacch
4	41	70.7	114	13	Q9Y42	Q9Y42	pantodon bu
5	40	69.0	242	4	Q9Y3T1	Q9Y3T1	homo sapien
6	40	69.0	275	16	Q9BBW0	Q9BBW0	rhizobium 1
7	40	69.0	349	16	Q9ZWH3	Q9ZWH3	rhizobium m
8	40	69.0	368	17	Q9Y7Y4	Q9Y7Y4	sulfolobus
9	40	69.0	477	2	Q93BD1	Q93BD1	mycobacteri-
10	40	69.0	962	5	Q9NGU5	Q9NGU5	droSOPhila
11	40	69.0	1006	5	Q9VTE2	Q9VTE2	droSOPhila
12	39	67.2	219	10	Q9FQE4	Q9FQE4	glycine max
13	39	67.2	224	10	Q9FQD8	Q9FQD8	glycine max
14	39	67.2	444	14	Q9RX02	Q9RX02	deinococcus
15	39	67.2	796	5	Q9VAP3	Q9VAP3	droSOPhila
16	38	65.5	158	11	Q9CYE7	Q9CYE7	mus musculus

## ALIGNMENTS

## RESULT 1.

RESULT	Q90143	Q90143	PRELIMINARY;	PRT;	115 AA.
ID	Q90143	Q90143			
IC	Q90143	Q90143			
DT	01-DEC-2001	(TrEMBLrel. 19, Created)			
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence up			
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation			
DE	PREPROSOMATOSTATIN.				
OS	Osteoglossum bicirrhosum (silver arawana).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Verte				
OC	Actinopterygii; Neopterygii; Teleostei; Oste				
OC	Osteoglossiformes; Osteoglossidae; Osteoglos				
OX	NCBI_Taxid=109271;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Al-Mahrouki A.A., Irwin D.M., Youson J.H.;				
RT	"Characterization of variant somatostatin CD				
RT	osteoglossomorphs: molecular identification				
RT	analysis."				
RL	Submitted (AUG-2000) to the EMBL/GenBank/DB				
RL	EMBL; AF292650; AAK97067.1;				
DR	EMBL; AF292650; AAK97067.1;				
SQ	SEQUENCE 115 AA; 12791 MW; D65F9BD7C6F1E4				

Query Match 79.3%; Score 46; DB 13; Length 115;  
Best Local Similarity 80.0%; Pred. No. 0.59;

QY 2 VDNLP PRK 11  
: : : : :  
Db 92 LNNLP PRK 101

RECITE

RESULT	2
Q9V466	PRELIMINARY; PRT; 845 AA.
ID	Q9V466
AC	Q9V466;
DT	01-MAY-2000 (TRENBLREL. 13, Created)
DT	01-MAY-2000 (TRENBLREL. 13, Last sequence update)



01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 BCDNA:LD18761 OR CG6743.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN-BERKELEY;  
 RC MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celnik S.E., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Amanatides P.G., Scher S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.Y.,  
 RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Herman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hostin D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matten B., McIntosh T.C., McLeod M.P., McPherson P.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mosurell A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 [2]  
 "The genome sequence of Drosophila melanogaster."  
 Science 287:2185-2195(2000).  
 [3]  
 SEQUENCE FROM N.A.  
 RP Tsang G., Brokstein P., Frise E., Harvey D., Evans-Holm M.,  
 RA Lewis S.E., Suh C., Rubin G.M.;  
 "Full length Drosophila melanogaster cDNA sequence."  
 Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF003629; AAF53002.1;  
 DR EMBL: AF160938; AAD46878.1;  
 DR FlyBase; Fggn0027868; BCDNA:LD18761.  
 SQ SEQUENCE 845 AA; 97361 MW; 0A75B807A6AFD6B6 CRC64;

Query Match 79.3%; Score 46; DB 5; Length 845;  
 Best Local Similarity 88.9%; Pred. No. 4.5;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVNDLPPE 9  
 |||||  
 Db 641 SVNDIPPE 649

RESULT 3  
 013787

ID 013787 PRELIMINARY; PRT; 273 AA.  
 AC 013787;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE PUTATIVE TRANSLOCATION PROTEIN C17G6.09.  
 GN SPAC17G6.09.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC NCBI\_TaxID=4896;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-972;  
 RC Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;  
 RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RL FUNCTION: REQUIRED FOR PREPROTEIN TRANSLOCATION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
 CC RETICULUM (BY SIMILARITY).  
 CC -1- SIMILARITY: TO YEAST AND Y. LIPOLYTICA SEC62.  
 DR EMBL: Z99162; CAB16220.1;  
 KW Hypothetical protein; Protein transport; Translocation; Transmembrane;  
 KW Endoplasmic reticulum.  
 FT DOMAIN 1 132 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 133 153 POTENTIAL.  
 FT DOMAIN 154 167 LUMENAL (POTENTIAL).  
 FT TRANSMEM 168 210 POTENTIAL.  
 FT DOMAIN 211 273 CYTOPLASMIC (POTENTIAL).  
 FT SEQUENCE 273 AA; 31585 MW; EE7A6369530799B3 CRC64;  
 SQ  
 Query Match 72.4%; Score 42; DB 3; Length 273;  
 Best Local Similarity 70.0%; Pred. No. 7.7;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VDNLPPEPK 11  
 |||||  
 DB 93 VDKLPKPKRK 102  
 RESULT 4  
 Q90Y42 PRELIMINARY; PRT; 114 AA.  
 ID Q90Y42;  
 AC Q90Y42;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE PREPROSOMATOSTATIN.  
 OS Pantodon buchholzi (Butterflyfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
 OC Osteoglossiformes; Pantodontidae; Pantodon.  
 OC NCBI\_TaxID=8276;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP Al-Nahrouki A.A., Irwin D.M., Youson J.H.;  
 RA "Characterization of variant somatostatin cDNAs from several  
 osteoglossomorphs: molecular identification and comparative  
 analysis."  
 Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF222651; AAK97068.1;  
 SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 70.7%; Score 41; DB 13; Length 114;  
 Best Local Similarity 87.5%; Pred. No. 4.8;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPPEPK 11  
 |||||  
 DB 93 NVPPPEPK 100

RESULT 5  
Q9Y3T1 PRELIMINARY; PRT; 242 AA.  
ID Q9Y3T1  
AC Q9Y3T1  
DT 01-NOV-1999 (TREMblrel. 12, Created)  
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE HYPOTHETICAL 28.3 KDA PROTEIN (FRAGMENT).  
GN DKFZP564A063.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eumetazoa; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Wambutt R., Heubner D., Newes H.W., Gassenhuber J., Wiemann S.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
JR EMBL; AL050008; CAB43235.1;  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 242 AA; 28250 MW; 84DFC201625C3B1F CRC64;

Query Match 69.0%; Score 40; DB 4; Length 242;  
Best Local Similarity 72.7%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11  
| | | | | | | |  
Db 145 SWDSLPPSERK 155

RESULT 6  
Q98BW0 PRELIMINARY; PRT; 275 AA.  
ID Q98BW0  
AC Q98BW0  
DT 01-OCT-2001 (TREMblrel. 18, Created)  
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DE MRS406 PROTEIN.  
GN MRS406  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MAFF303099;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Rhizobium loti".  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003006; BAB51862.1;  
DR InterPro; IPR001763; Rhodanese\_domain.  
DR SMART; SM00450; RHOD; 1.  
KW Complete proteome.  
SQ SEQUENCE 275 AA; 30613 MW; E12618905C17D6E5 CRC64;

Query Match 69.0%; Score 40; DB 16; Length 275;  
Best Local Similarity 63.6%; Pred. No. 18;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11  
| | | | | | | |  
Db 110 STDKLPPEPRK 120

RESULT 7  
Q92WH3 PRELIMINARY; PRT; 349 AA.  
ID Q92WH3  
AC Q92WH3  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE PUTATIVE ABC TRANSPORTER ATP-BINDING PROTEIN.  
GN SMG20360.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Plasmid pSymb (megaplasmid 2).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1021;  
RX MEDLINE-21396508; PubMed-11481431;  
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,  
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Gouzy J.,  
RA Golding B., Puehler A.;  
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-  
fixing endosymbiont Sinorhizobium meliloti".  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
DR EMBL; AL603643; CAC48766.1;  
KW ATP-binding; Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 349 AA; 38734 MW; 17DCB997F3646393 CRC64;

Query Match 69.0%; Score 40; DB 16; Length 349;  
Best Local Similarity 60.0%; Pred. No. 23;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 10  
| | | | | | | |  
Db 69 AIDNLPPEPRK 78

RESULT 8  
Q97YY4 PRELIMINARY; PRT; 368 AA.  
ID Q97YY4  
AC Q97YY4  
DT 01-OCT-2001 (TREMblrel. 18, Created)  
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DE SUGAR ABC TRANSPORTER.  
GN SS01168.  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=2287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE-21332296; PubMed-11427726;  
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
RA Aveyez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
RA De Moers A., Etraus G., Fletcher C., Gordon P.M.K.,  
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2".  
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
DR EMBL; AE006733; AAK41416.1;  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR003439; ABC\_transporter.  
DR InterPro; IPR001687; ATP\_GTP\_A.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR SMART; SM00382; AAA; 1.  
KW Complete proteome.  
SQ SEQUENCE 368 AA; 41884 MW; 6B728206043290F5 CRC64;

Query Match 69.0%; Score 40; DB 17; Length 368;  
 Best Local Similarity 60.0%; Pred. No. 24;  
 Matches 6; Conservative 4; Mismatches 0; Gaps 0;

Qy 2 VDNLPKRR 11  
 Db 72 VDDLPKRR 81

RESULT 9  
 ID Q938DI PRELIMINARY; PRT; 477 AA.

AC Q938DI; 2001 (TREMELrel. 19, Created)  
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
 DE HYPOTHETICAL 50.9 KDA PROTEIN.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1772;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MC2.155;  
 RA Derbyshire K.M., Parsons L.M., DeVost J.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY054120; AAL17929.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 477 AA; 50879 MW; 6EC7788CA2CDE913 CRC64;

Query Match 69.0%; Score 40; DB 2; Length 477;  
 Best Local Similarity 70.0%; Pred. No. 32;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SVDNLPKRR 10  
 Db 384 SVDNLPKRR 393

RESULT 10  
 ID Q9N6U5 PRELIMINARY; PRT; 962 AA.

AC Q9N6U5; 2000 (TREMELrel. 15, Created)  
 DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)  
 DE SU(UR)ES OR CG7869.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RP SEQUENCE FROM N.A.  
 RC Makunin I.V.;  
 RT "Characterization of the Su(UR)ES gene."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ277592; CAB89187.1;  
 DR FlyBase; FBgn0025355; Su(UR)ES.  
 DR InterPro; IPR000330; SNF2\_N.  
 DR Pfam; PF00176; SNF2\_N; 1.  
 SQ SEQUENCE 962 AA; 107624 MW; 5063FB88E875C4F9 CRC64;

Query Match 69.0%; Score 40; DB 5; Length 962;  
 Best Local Similarity 60.0%; Pred. No. 65;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SVDNLPKRR 10

Db 400 ATDNMPKRR 409

RESULT 11  
 ID Q9VTE2 PRELIMINARY; PRT; 1006 AA.

AC Q9VTE2; 2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2001 (TREMELrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)  
 DE SUPPRESSOR OF UNDERREPLICATION ES-PROTEIN.  
 GN SU(UR)ES OR CG7869.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Arif J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Bensch P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablo B., DeCher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Fodor C., Gabriellian A.C., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,  
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs K.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003546; AAF50110.2;  
 DR FlyBase; FBgn0025355; Su(UR)ES.  
 DR InterPro; IPR000330; SNF2\_N.  
 DR Pfam; PF00176; SNF2\_N; 1.  
 SQ SEQUENCE 1006 AA; 112638 MW; C281A9520DD242A CRC64;

Query Match 69.0%; Score 40; DB 5; Length 1006;  
 Best Local Similarity 60.0%; Pred. No. 68;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SVDNLPKRR 10

DB 444 ATDNPPRKR 453

RESULT 12  
Q9FOE4 PRELIMINARY; PRT: 219 AA.  
AC Q9FOE4; 2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DE GLUTATHIONE S-TRANSFERASE GST 14 (EC 2.5.1.18) (FRAGMENT).  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20532786; PubMed=11080288;  
RT "A Genomics Approach to the Comprehensive Analysis of the Glutathione  
S-Transferase Gene Family in Soybean and Maize."  
RL Plant Physiol. 124:1105-1120(2000).  
DR EMBL: AF243369; AAG34804.1;  
DR InterPro: IPR004046; GST\_C;  
DR InterPro: IPR004045; GST\_N.  
DR Pfam: PF00043; GST\_C; 1.  
DR Pfam: PF00043; GST\_N; 1.  
KW transferase.  
FT NON\_TER  
SQ SEQUENCE 219 AA; 25356 MW; 7378601EFB6C7630 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 219;  
Best Local Similarity 75.0%; Pred. No. 22;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 DNLPPR 10  
DB 194 DNLPPR 201

RESULT 13  
Q9FQD8 PRELIMINARY; PRT: 224 AA.  
AC Q9FQD8; 2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DE GLUTATHIONE S-TRANSFERASE GST 20 (EC 2.5.1.18).  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20532786; PubMed=11080288;  
RT "A Genomics Approach to the Comprehensive Analysis of the Glutathione  
S-Transferase Gene Family in Soybean and Maize."  
RL Plant Physiol. 124:1105-1120(2000).  
DR EMBL: AF243375; AAG34810.1;  
DR InterPro: IPR004046; GST\_C;  
DR InterPro: IPR004045; GST\_N.  
DR Pfam: PF00043; GST\_C; 1.  
DR Pfam: PF02798; GST\_N; 1.  
KW transferase.  
FT NON\_TER  
SQ SEQUENCE 224 AA; 26222 MW; 013C0D0C72D630DF CRC64;

Query Match 67.2%; Score 39; DB 10; Length 224;  
Best Local Similarity 75.0%; Pred. No. 22;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 DNLPPR 10  
DB 202 DNLPPR 209

RESULT 14  
Q9RX02 PRELIMINARY; PRT: 444 AA.  
AC Q9RX02; 2001 (TRENBLREL. 13, Created)  
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
DE 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
OS INTEGRASE, PUTATIVE.  
GN DR0513.  
OC Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20036896; PubMed=10567266;  
RT White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
Fraser C.M.;  
RT radiodurans RI.  
RL Science 286:1571-1577(1999).  
DR EMBL: AE001910; AAF10092.1;  
DR TIGR: DR0513;  
DR InterPro: IPR001994; Cytidylyltransf.  
DR InterPro: IPR002104; Phage\_integrase.  
DR Pfam: PF00589; Phage\_integrase; 1.  
KW Complete proteome.  
SQ SEQUENCE 444 AA; 48293 MW; A885775BD76F0143 CRC64;

Query Match 67.2%; Score 39; DB 16; Length 444;  
Best Local Similarity 87.5%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 DNLPPR 10  
DB 67 DNLPPR 74

RESULT 15  
Q9VAP3 PRELIMINARY; PRT: 796 AA.  
AC Q9VAP3; 2001 (TRENBLREL. 13, Created)  
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
DE 01-JUN-2001 (TRENBLREL. 17, Last annotation update)  
GN CG11880-PROTEIN.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RT Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu E., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center H., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster".  
 RL Science 287:2185-2193(2000).  
 DR EMBL: AE003768; AAF56859.1; -  
 DR FlyBase; FBgn0039637; CG11880.  
 SQ SEQUENCE 796 AA; 90628 MW; 20961DD889A3EE3B CRC64;

Query Match 67.2%; Score 39; DB 5; Length 796;  
 Best Local Similarity 66.7%; Pred. No. 82;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRRK 11  
 :|||||:  
 Db 786 NNLPPRRR 794

Search completed: June 13, 2002, 12:31:03  
 Db time: 485 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:29:54 ; Search time 18.73 Seconds  
(without alignments)  
22.740 Million cell updates/sec

Title: US-09-727-739b-18  
Perfect score: 58  
Sequence: 1 SVDNLPPEPRK 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	84.5	125	1	SMS2_LOPAM
2	45	77.6	73	1	SMS2_PLAFPE
3	45	77.6	115	1	SMS2_ONCMY
4	40	69.0	1061	1	TRC4_ECOLI
5	39	67.2	110	1	YFMT_TETH
6	38	65.5	28	1	SMS2_ORENI
7	37	63.8	76	1	SIFA_ARATH
8	37	63.8	78	1	Y321_HAEIN
9	37	63.8	124	1	YIDD_STRCO
10	37	63.8	250	1	RL7B_SCHPO
11	37	63.8	251	1	RL7C_SCHPO
12	37	63.8	339	1	SYFA_DEIRA
13	37	63.8	424	1	PSGC_HUMAN
14	37	63.8	424	1	PSGC_HUMAN
15	37	63.8	435	1	PSG6_HUMAN
16	37	63.8	641	1	HAPK_PSESY
17	37	63.8	1955	1	PC15_HUMAN
18	36	62.1	246	1	BRM1_HUMAN
19	36	62.1	335	1	MAUG_METME
20	36	62.1	392	1	ORIA_DROME
21	36	62.1	431	1	VGF_BPAL3
22	36	62.1	431	1	VGF_BPAPH
23	36	62.1	458	1	YNEL_CAEEL
24	36	62.1	1343	1	PC15_MOUSE
25	35	60.3	325	1	SYFA_THEMEA
26	35	60.3	345	1	TF2B_YEAST
27	35	60.3	359	1	C11D_YEAST
28	35	60.3	406	1	CG16_HUMAN
29	35	60.3	419	1	PSG1_HUMAN
30	35	60.3	419	1	PSG4_HUMAN
31	35	60.3	419	1	PSG7_HUMAN
32	35	60.3	426	1	PSG8_HUMAN
33	35	60.3	428	1	PSG3_HUMAN

34	35	60.3	495	1	OXAL_HUMAN
35	35	60.3	591	1	PAKA_HUMAN
36	35	60.3	714	1	YFL2_YEAST
37	35	60.3	825	1	AFP_SCHPO
38	35	60.3	1221	1	TOP2_TRIBB
39	35	60.3	1506	1	PK3G_MOUSE
40	34.5	59.5	74	1	SMS2_MYOSC
41	34	58.6	70	1	SIFA_SPIOL
42	34	58.6	155	1	HS21_PHANI
43	34	58.6	222	1	Y066_RICPR
44	34	58.6	249	1	MCRG_METFE
45	34	58.6	249	1	RL7A_SCHPO

## ALIGNMENTS

RESULT 1  
SMS2\_LOPAM STANDARD; PRT; 125 AA.  
AC P01170: Q91066;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr7,Gly10]somatostatin-14].  
OS Lophius americanus (American goosefish) (Anglerfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.  
OX NCBI\_TaxID=8073;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81052423; PubMed=6107860;  
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;  
RT "Cloning and sequence analysis of cDNAs encoding two distinct somatostatin precursors found in the endocrine pancreas of anglerfish";  
FT anglerfish";  
RL Nature 288:137-141(1980).  
RN [2]  
RP PARTIAL SEQUENCE, AND HYDROXYLATION.  
RX MEDLINE=87308304; PubMed=2887572;  
RA Andrews P.C., Nichols R., Dixon J.E.;  
RT "Post-translational processing of preprosomatostatin-II examined using fast atom bombardment mass spectrometry";  
RL J. Biol. Chem. 262:12692-12699(1987).  
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.  
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; V00641; CAA23987.1;  
CC PIR; A01434; RIAFS2.  
DR PIR; A27376; A27376.  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;  
KW Multigene family.  
FT SIGNAL 1 24 POTENTIAL.  
FT PROPEP 25 109  
FT PEPTIDE 112 125 [TYR7, GLY10]SOMATOSTATIN-14.  
FT DISULFID 114 125  
FT MOD\_RES 120 120 HYDROXYLATION  
FT CONFLICT 77 78 DV->TG (IN REF. 1).  
FT CONFLICT 90 90 G->E (IN REF. 1).

SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 84.5%; Score 49; DB 1; Length 125;  
Best Local Similarity 81.8%; Pred. No. 0.027; 1; Indels 0; Gaps 0;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SVDNLPPEK 11  
| : |||||  
DB 101 STNNLPPEK 111

RESULT 2  
SMS2\_PLAFE STANDARD; PRT; 73 AA.  
AC P21780;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;  
[Tyr7,Gly10]somatostatin-14] (Fragments)  
E Placichthys flesus (European flounder);  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
CC Acanthopterygia; Acanthopterygii; Percomorpha; Pleuronectiformes;  
CC Pleuronectoidae; Pleuronectidae; Placichthys.  
OX NCBI\_TaxID=8260;  
RN [1]  
RP SEQUENCE.

TX TISSUE=Pancreas;  
RX MEDLINE=88029486; PubMed=2889597;  
RA Conlon J.M., Davis M.S., Falkmer S., Thim L.;  
RT "Structural characterization of peptides derived from  
RT somatostatins I and II isolated from the pancreatic islets of two  
RT species of teleostean fish: the daddy sculpin and the flounder.";  
RL Eur. J. Biochem. 168:647-652(1987).  
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
DR PIR: S00169; S00169.  
KW Cleavage on pair of basic residues; Hormone; Multigene family.  
FT NON\_TER 1  
FT NON\_CONS 10 11  
FT PEPTIDE 45 46 [TYR21, GLY24]SOMATOSTATIN-28.  
FT PEPTIDE 46 73 [TYR7, GLY10]SOMATOSTATIN-14.  
FT PEPTIDE 60 73  
FT DISULFID 62 73  
SQ SEQUENCE 73 AA; 7989 MW; CCCBA6B30DCB29BB CRC64;

Query Match 77.6%; Score 45; DB 1; Length 73;  
Best Local Similarity 88.9%; Pred. No. 0.079; 0; Indels 0; Gaps 0;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 DNLPPPEK 11  
| : |||||  
DB 51 NNLPPEK 59

RESULT 3  
SMS2\_ONCMY STANDARD; PRT; 115 AA.  
AC Q91194;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;  
[Tyr7,Gly10]somatostatin-14].  
E Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;

[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=95354921; PubMed=7628684;  
RA Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;  
RT "Isolation and characterization of a cDNA encoding for  
RT preprosomatostatin containing [Tyr7, Gly10]-somatostatin-14 from the  
RT endocrine pancreas of rainbow trout, Oncorhynchus mykiss.";  
RL Gen. Comp. Endocrinol. 98:253-261(1995).  
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: U32471; AAC59695.1;  
DR InterPro: IPR004250; Somatostatin.  
DR Pfam: PF03002; Somatostatin\_1  
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 87  
FT PEPTIDE 88 115 [TYR21, GLY24]SOMATOSTATIN-28 (POTENTIAL).  
FT PEPTIDE 102 115 [TYR7, GLY10]SOMATOSTATIN-14.  
FT DISULFID 104 115 BY SIMILARITY.  
SQ SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;

Query Match 77.6%; Score 45; DB 1; Length 115;  
Best Local Similarity 88.9%; Pred. No. 0.13;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 DNLPPPEK 11  
| : |||||  
DB 93 NNLPPEK 101

RESULT 4  
TRC4\_ECOLI STANDARD; PRT; 1061 AA.  
ID TRC4\_ECOLI  
AC P27189; P27184;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE DNA primase traC (EC 2.7.7.-) (Replication primase).  
GN TRAC.  
OS Escherichia coli.  
OG Plasmid Incp-beta RP4.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14 AND 316-333.  
RX STRAIN=HB101.  
RX MEDLINE=92297959; PubMed=1818755;  
RT "Gene organization and nucleotide sequence of the primase region of  
RT Incp plasmids RP4 and R751.";  
RL DNA Seq. 2:1145-162(1991).  
CC -1- FUNCTION: REQUIRED FOR AUTONOMOUS REPLICATION IN ESCHERICHIA COLI.  
CC TRANSFERRED INTO THE RECIPIENT CELL DURING BACTERIAL CONJUGATION.  
CC CATALYZES THE SYNTHESIS OF SHORT OLIGONUCLEOTIDE PRIMERS WITH  
CC CPA... OR PCPA... AT THEIR 5'-TERMINI ON A SINGLE STRANDED  
CC TEMPLATE DNA.  
CC -1- ALTERNATIVE PRODUCTS: THE TRAC-1 AND TRAC-2 PROTEINS ARE PRODUCED  
CC BY THE USE OF ALTERNATIVE INITIATION SITES.  
CC -1- SIMILARITY: TO PLASMID R751 TRAC.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration





CC 1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC 1- SIMILARITY: STRONG, TO SIFA IN OTHER PLANTS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: T20524; ; NOT\_ANNOTATED\_CDS.  
 CC KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.  
 CC SEQUENCE 76 AA; 8252 MW; F05E3D595F8A9AF CRC64;

Query Match 63.8%; Score 37; DB 1; Length 76;  
 Best Local Similarity 75.0%; Pred. No. 2.5;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 4 NLPPRRK 11  
 DB 48 NLPPRRKK 55

RESULT 8  
 Y321\_HAEIN STANDARD; PRT; 78 AA.  
 ID Y321\_HAEIN  
 AC Q57534;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Protein H10321.  
 GN H10321.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
 OC Haemophilus.  
 OC NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clifton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghagan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.;  
 RL Science 269:496-512(1995).  
 RN [2]  
 RP IDENTIFICATION BY MASS SPECTROMETRY.  
 RX MEDLINE=20137488; PubMed=10675023;  
 RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,  
 RA Gray C., Fountoulakis M.;  
 RT "Two-dimensional map of the proteome of Haemophilus influenzae";  
 RL Electrophoresis 21:411-429(2000).  
 CC 1- SIMILARITY: SOME, TO B.NODOSUS VIRULENCE-ASSOCIATED PROTEIN B  
 CC (VAPB).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U32717; AAC21984.1;

DR TIGR: H10321; -;  
 KW Complete proteome.  
 SQ SEQUENCE 78 AA; 9037 MW; 468D048ED7C5C720 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 78;  
 Best Local Similarity 66.7%; Pred. No. 2.6;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DMLPPRRK 11  
 DB 68 DMLPPQERE 76

RESULT 9  
 YIDD\_STRCO STANDARD; PRT; 124 AA.  
 ID YIDD\_STRCO  
 AC O54625;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 13.6 kDa protein SH24.04.  
 GN SH24.04.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 OC NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=99014240; PubMed=9795152;  
 RA Gal-Mor O., Borovok I., Av-Gay Y., Cohen G., Aharonovitz Y.;  
 RA "Gene organization in the trxA/B-oric region of the Streptomyces  
 RT coelicolor chromosome and comparison with other eubacteria";  
 RL Gene 217:83-90(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Kim H.-D., Calcutt M.J., Schmidt F.J., Chater K.F.;  
 RT "Partitioning of the linear chromosome during sporulation of  
 RT Streptomyces coelicolor A3(2) involves an ORC-linked parB locus.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Oliver K., Harris D., James K.D., Parkhill J., Barrell B.G.,  
 RA Rajandream M.A.;  
 RA Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RL 1- SIMILARITY: BELONGS TO THE UPF0161 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Y16311; CA476173.1;  
 DR EMBL: AL049826; CAB42698.1;  
 DR EMBL: AF187159; AAF16009.1;  
 DR EMBL: AF031590; AAC03488.1;  
 DR InterPro: IPR002696; DUF37.  
 DR Pfam: PF01809; DUF37; 1.  
 DR ProDom: PD004225; DUF37; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 124 AA; 13599 MW; 8BB3AD1786D8075E CRC64;

Query Match 63.8%; Score 37; DB 1; Length 124;  
 Best Local Similarity 66.7%; Pred. No. 4.4;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPPIPRK 10  
|:|:|:|:|  
DB 65 VDHVPPRK 73

RESULT 10  
RL7B\_SCHPO STANDARD; PRT: 250 AA.  
AC P25457; RL7B\_SCHPO  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE 60S ribosomal protein L7-B.  
GN RPL7B OR SPAC3H5.07.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
NCBI\_TaxID=4896;  
X [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=91212191; PubMed=2020547;  
RA Damagene V., Recondo A., Baldacci G.;  
RT Identification of a gene encoding the predicted ribosomal protein  
RT L7b divergently transcribed from POL1 in fission yeast  
PT Schizosaccharomycetes pombe".  
PL Nucleic Acids Res. 19:1099-1104(1991).  
RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN=972;  
RA Oliver K., Harris D., Wood V., Barrell B.G., Rajandream M.A.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- MISCELLANEOUS: THERE ARE THREE GENES FOR L7 IN S.POMBE.  
CC -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X54983; CAA38729.1;  
DR EMBL; 299296; CAB16592.1;  
DR PIR; S25067; S25067.  
OR InterPro: IPR000517; Ribosomal\_L30.  
R Pfam: PF00327; Ribosomal\_L30; 1.  
OR PROSITE; PS00634; RIBOSOMAL\_L30; 1.  
KW Ribosomal protein; Multigene family.  
SQ SEQUENCE 250 AA; 28449 MW; 2AE917F0DAF7A8CB CRC64;

Query Match 63.8%; Score 37; DB 1; Length 250;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPPIPRK 11  
|:|:|:|:|  
DB 99 INNIPPKARK 108

RESULT 11  
ID RL7C\_SCHPO STANDARD; PRT: 251 AA.  
AC O60143;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE 60S ribosomal protein L7-C.  
GN RPL7C OR SPCL18H10.12C.  
OS Schizosaccharomycetes pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;  
RA Lyne M., Rajandream M.A., Barrell B.G., Badcock K., Churcher C.M.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- MISCELLANEOUS: THERE ARE THREE GENES FOR L7 IN S.POMBE.

CC -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AL022304; CAA18409.1;  
DR InterPro: IPR000517; Ribosomal\_L30.  
DR Pfam: PF00327; Ribosomal\_L30; 1.  
DR PROSITE; PS00634; RIBOSOMAL\_L30; 1.  
KW Ribosomal protein; Multigene family.  
SQ SEQUENCE 251 AA; 28730 MW; 5382A27E39F3800E CRC64;

Query Match 63.8%; Score 37; DB 1; Length 251;  
Best Local Similarity 50.0%; Pred. No. 10;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPPIPRK 11  
|:|:|:|:|  
DB 100 INNIPPKARK 109

RESULT 12

SYFA\_DEIRA

ID SYFA\_DEIRA STANDARD; PRT: 339 AA.

AC Q9RRX8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20) (Phenylalanine-

tRNA ligase alpha chain) (PheRS).

GN PHE5 OR DR2354

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.

OX NCBI\_TaxID=1299;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RL;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans RL."

RL Science 286:1571-1577(1999).

CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +

CC diphosphate + L-phenylalanyl-tRNA(Phe).

CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.

CC PHE-TRNA SYNTHETASE ALPHA CHAIN SUBFAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



```

CC DR EMBL; AC005260; AAC25620.1;
DR MIM; 176399;
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00047; Ig; 4;
DR SMART; SM00410; Ig_Like; 3.
DR SMART; SM00408; IGC2; 1.
KW Immunoglobulin domain; Glycoprotein; Signal; Repeat; Multigene family.
FT SIGNAL 1 34
FT CHAIN 35 424
FT DOMAIN 35 143
FT DOMAIN 161 223
FT DOMAIN 254 316
FT DOMAIN 346 400
FT SITE 126 128
FT SITE 168 216
FT DISULFID 261 309
FT DISULFID 261 309
FT CARBOHYD 103 103
FT CARBOHYD 110 110
FT CARBOHYD 198 198
FT CARBOHYD 267 267
FT CARBOHYD 302 302
FT CARBOHYD 386 386
FT CARBOHYD 419 419
SQ SEQUENCE 424 AA; 47501 MW; 6ECB345E53CCDEA4 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 424;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11
DB 244 TNNLNPEPRK 254

RESULT 15
ID PSG6_HUMAN STANDARD; PRT; 435 AA.
AC Q00889;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
JE Pregnancy-specific beta-1-glycoprotein 6 precursor (PSBG-6).
GN PSG6 OR CM3 OR PSGGB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC [1]
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-89352023; PubMed-2783133;
RA Zimmermann W., Weiss M., Thompson J.A.;
RT cDNA cloning demonstrates the expression of pregnancy-specific
RT glycoprotein genes, a subgroup of the carcinoembryonic antigen gene
RT family, in fetal liver.;
RL Biochem. Biophys. Res. Commun. 163:1197-1209(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-90211271; PubMed-1690992;
RA Thompson J., Kourami R., Wagner K., Barnert S., Schleussner C.,
RA Schrewe H., Zimmermann W., Mueller G., Schempp W., Zaninetta D.,
RA Ammaturo D., Hardman N.;
RT "The human pregnancy-specific glycoprotein genes are tightly linked on
RT the long arm of chromosome 19 and are coordinately expressed.";
RL Biochem. Biophys. Res. Commun. 167:848-859(1990).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RA Lanerding J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do D., Regalia W., Terry A., Ganes J.,
RA Danganan L., Poundstone F., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 23-142 FROM N.A.
RX MEDLINE-89149826; PubMed-2537643;
RA Thompson J.A., Mauch E.M., Chen F.S., Hinoda Y., Schrewe H.,
RA Belling B., Barnert S., von Kleist S., Shively J.E., Zimmermann W.;
RT "Analysis of the size of the carcinoembryonic antigen (CEA) gene
RT family: Isolation and sequencing of N-terminal domain exons.";
RL Biochem. Biophys. Res. Commun. 158:996-1004(1989).
RN [5]
RP SEQUENCE OF 1-142 FROM N.A.
RX MEDLINE-9032674; PubMed-2377620;
RA Leslie K.K., Watanabe S., Lei K.-J., Chou D.Y., Plouzek C.A.,
RA Deng H.-C., Torres J., Chou J.F.;
RT "Linkage of two human pregnancy-specific beta 1-glycoprotein genes:
RT one is associated with hydatidiform mole.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5822-5826(1990).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- DEVELOPMENTAL STAGE: PSBG ARE PRODUCED IN HIGH QUANTITY DURING
CC PREGNANCY.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M33666; AAC60206.1;
DR EMBL; M32621; AAC60201.1;
DR EMBL; M32620; AAC60201.1; JOINED.
DR EMBL; AC005260; AAC25619.1;
DR EMBL; M37757; AAC60202.1;
DR EMBL; M37756; AAC60202.1; JOINED.
DR MIN; 176395;
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00410; Ig_Like; 3.
DR SMART; SM00408; IGC2; 1.
KW Polymorphism.
FT SIGNAL 1 34
FT CHAIN 35 435
FT DOMAIN 161 223
FT DOMAIN 254 316
FT DOMAIN 346 400
FT SITE 126 128
FT DISULFID 168 216
FT DISULFID 261 309
FT DISULFID 353 393
FT CARBOHYD 103 103
FT CARBOHYD 110 110
FT CARBOHYD 198 198
FT CARBOHYD 267 267
FT CARBOHYD 302 302
FT CARBOHYD 386 386
FT CARBOHYD 419 419

```

```

FT CARBOHYD 386 386 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 155 155 L -> F (IN DBSNP:1058674).
FT 165 165 /FTIG-VAR_011723.
FT VARIANT 167 167 R -> L (IN DBSNP:1058680).
FT 180 180 /FTIG-VAR_011724.
FT 181 181 I -> T (IN DBSNP:1065505).
FT 185 185 /FTIG-VAR_011725.
FT 191 191 L -> W (IN DBSNP:1065507).
FT 196 196 /FTIG-VAR_011726.
FT 253 253 L -> M (IN DBSNP:1065508).
FT 257 257 /FTIG-VAR_011727.
FT 258 258 N -> S (IN DBSNP:1065509).
FT 344 344 /FTIG-VAR_011728.
FT 404 404 R -> S (IN DBSNP:1058688).
FT 435 435 /FTIG-VAR_011729.
FT 48813 48813 MW, D94D4B0884327C60 CRC64;
SQ SEQUENCE 435 AA; 48813 MW, D94D4B0884327C60 CRC64;

```

Query Match 63.8%; Score 37; DB 1; Length 435;  
 Best Local Similarity 54.5%; Pred. No. 19;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 SVDNLPPEKK 11  
 DB 244 TNNLPPEKK 254

Search completed: June 13, 2002, 12:29:54  
 Job time: 446 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:56 ; Search time 28.68 Seconds  
(without alignments)  
9.368 Million cell updates/sec

Title: US-09-727-739B-18  
Perfect score: 58  
Sequence: 1 SYDNLPFRERK 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_5/ptodata/2/1aa/5A-COMB.pep.\*  
2: /cgn2\_5/ptodata/2/1aa/5B-COMB.pep.\*  
3: /cgn2\_5/ptodata/2/1aa/6A-COMB.pep.\*  
4: /cgn2\_5/ptodata/2/1aa/6B-COMB.pep.\*  
5: /cgn2\_5/ptodata/2/1aa/PTCUS-COMB.pep.\*  
6: /cgn2\_5/ptodata/2/1aa/BACKFILES.pep.\*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	39	67.2	219	4	US-09-247-373B-52
2	39	67.2	224	4	US-09-247-373B-34
3	37	63.8	424	6	5169835-6
4	36	62.1	2496	4	US-09-125-028-2
5	36	62.1	2958	4	US-08-894-344C-2
6	35	60.3	230	6	5169835-13
7	35	60.3	419	6	5169835-2
8	35	60.3	591	3	US-09-082-737-2
9	34	58.6	223	3	US-08-857-534-12
10	34	58.6	223	5	PCR-US95-04971-12
11	34	58.6	229	4	US-09-247-373B-48
12	34	58.6	361	1	US-08-258-261B-3
13	34	58.6	361	1	US-08-456-837-3
14	34	58.6	361	1	US-08-457-342-3
15	34	58.6	361	1	US-08-457-646A-3
16	34	58.6	361	1	US-08-458-076A-3
17	34	58.6	361	1	US-08-457-335A-3
18	34	58.6	361	2	US-08-729-214-3
19	34	58.6	361	3	US-09-028-934-3
20	33	56.9	265	2	US-08-970-133-1
21	33	56.9	265	4	US-09-294-545-1
22	33	56.9	345	1	US-08-031-148-4
23	33	56.9	345	2	US-08-415-838-4
24	33	56.9	415	2	US-08-576-626A-52
25	33	56.9	683	1	US-07-878-960-2
26	33	56.9	855	1	US-08-344-536-2
27	33	56.9	855	1	US-08-344-536-5

28	33	56.9	855	3	US-08-920-562-2	Sequence 2, Appl1
29	33	56.9	855	3	US-08-920-562-5	Sequence 5, Appl1
30	33	56.9	2544	2	US-08-576-626A-32	Sequence 32, Appl1
31	32	55.2	275	4	US-09-147-915-4	Sequence 4, Appl1
32	32	55.2	296	3	US-08-986-769-2	Sequence 2, Appl1
33	32	55.2	361	3	US-09-028-934-30	Sequence 30, Appl1
34	32	55.2	449	1	US-08-624-663A-2	Sequence 2, Appl1
35	32	55.2	449	2	US-08-974-565C-1	Sequence 1, Appl1
36	32	55.2	449	3	US-09-255-748-1	Sequence 1, Appl1
37	32	55.2	470	4	US-09-004-838-44	Sequence 44, Appl1
38	32	55.2	472	4	US-09-004-838-43	Sequence 43, Appl1
39	32	55.2	560	2	US-08-643-034A-2	Sequence 2, Appl1
40	32	55.2	560	3	US-08-648-650A-2	Sequence 2, Appl1
41	32	55.2	600	3	US-08-904-871-3	Sequence 3, Appl1
42	32	55.2	713	2	US-08-987-466-3	Sequence 3, Appl1
43	32	55.2	713	3	US-08-974-565C-5	Sequence 5, Appl1
44	32	55.2	713	3	US-09-255-748-5	Sequence 5, Appl1
45	32	55.2	713	4	US-09-240-359-3	Sequence 3, Appl1

## ALIGNMENTS

RESULT 1  
US-09-247-373B-52  
; Sequence 52, Application US/09247373B  
; Patent No. 6168954  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CI-1108-A  
; CURRENT APPLICATION NUMBER: US/09/247,373B  
; CURRENT FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 08/924,747  
; PRIOR FILING DATE: 1997-09-05  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 52  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: SOYBEAN  
US-09-247-373B-52

Query Match 67.2% Score 39; DB 4; Length 219;  
Best Local Similarity 75.0% Pred. No. 8.5;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 DNLPPRR 10  
DB 194 DNLPPRDK 201

RESULT 2  
US-09-247-373B-34  
; Sequence 34, Application US/09247373B  
; Patent No. 6168954  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CI-1108-A  
; CURRENT APPLICATION NUMBER: US/09/247,373B  
; CURRENT FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 08/924,747  
; PRIOR FILING DATE: 1997-09-05  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 34  
; LENGTH: 224  
; TYPE: PRT  
; ORGANISM: SOYBEAN

US-09-247-373B-34

Query Match 67.2%; Score 39; DB 4; Length 224;  
Best Local Similarity 75.0%; Pred. No. 8.7;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10  
DB 202 DNLPPRRK 209

RESULT 3

5169835-6  
; Patent No. 5169835  
; APPLICANT: WAI-YEE, CHAN  
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS  
; NUMBER OF SEQUENCES: 48  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-AUG-1989  
; SEQ ID NO: 6  
; LENGTH: 424

5169835-6

Query Match 63.8%; Score 37; DB 6; Length 424;  
Best Local Similarity 54.5%; Pred. No. 39;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPERRK 11  
DB 244 TNNLPERRK 254

RESULT 4

US-09-125-028-2  
; Sequence 2, Application US/09125028A  
; Patent No. 6190707  
; GENERAL INFORMATION:  
; APPLICANT: WADOUX, Isabelle  
; APPLICANT: COLAVIZZA, Didier  
; APPLICANT: LOIEZ, Annie  
; TITLE OF INVENTION: NOVEL COLD-SENSITIVE BREAD-MAKING YEASTS.  
; FILE REFERENCE: levure sensible froid  
; CURRENT APPLICATION NUMBER: US/09/125,028A  
; EARLIER FILING DATE: 1998-08-07  
; EARLIER FILING DATE: 1997-02-07  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2496  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae

US-09-125-028-2

Query Match 62.1%; Score 36; DB 4; Length 2496;  
Best Local Similarity 66.7%; Pred. No. 3.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPERR 10  
DB 257 IDVLPERRK 265

RESULT 5

US-08-894-344C-2  
; Sequence 2, Application US/08894344C  
; Patent No. 6172196  
; GENERAL INFORMATION:  
; APPLICANT: KAWASAKI, Hideki

APPLICANT: TOKAI, Masaya  
APPLICANT: KIKUCHI, Yasuhiro  
APPLICANT: OUCHI, KOZO  
TITLE OF INVENTION: DNA ENCODING PROTEIN COMPLEMENTING  
TITLE OF INVENTION: YEAST  
TITLE OF INVENTION: LOW TEMPERATURE-SENSITIVE FERMENTABILITY  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112-3601

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 1440 Kb storage.

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/894,344C

FILING DATE: 15-AUGUST-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP343700/95

FILING DATE: 28-DECEMBER-1995

APPLICATION NUMBER: PCT/JP96/03862

FILING DATE: 27-DECEMBER-1996

ATTORNEY/AGENT INFORMATION:

NAME: Perry, Lawrence S.

REGISTRATION NUMBER: 31865

TELEPHONE: 212-218-2100

TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2958 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Saccharomyces cerevisiae

STRAIN: X2180-1B

US-08-894-344C-2

Query Match 62.1%; Score 36; DB 4; Length 2958;

Best Local Similarity 66.7%; Pred. No. 4.6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VDNLPERR 10

DB 257 IDVLPERRK 265

RESULT 6

5169835-13  
; Patent No. 5169835  
; APPLICANT: WAI-YEE, CHAN  
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS  
; NUMBER OF SEQUENCES: 48  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-AUG-1989  
; SEQ ID NO: 13  
; LENGTH: 230

5169835-13

Query Match 60.3%; Score 35; DB 6; Length 230;  
Best Local Similarity 54.5%; Pred. No. 46;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPERRK 11





COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04971  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/235,406  
FILING DATE: APRIL 28, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Luann Cseri  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-004/00WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-843-5165  
TELEFAX: 415-857-0663  
TELEX: 380816 Coolcypa  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 223 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
PCT-US95-04971-12

Query Match 58.6%; Score 34; DB 5; Length 223;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPRERK 11  
Db 44 PPRERK 49

RESULT 11  
US-09-247-373b-48  
Sequence 48; Application US/09247373B  
Patent No. 6168954  
GENERAL INFORMATION:  
APPLICANT: MCGONIGLE, BRIAN  
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
FILE REFERENCE: CL-1108-A  
CURRENT APPLICATION NUMBER: US/09/247,373B  
CURRENT FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 08/924,747  
PRIOR FILING DATE: 1997-09-05  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 48  
LENGTH: 229  
TYPE: PRT  
ORGANISM: SOYBEAN  
US-09-247-373b-48

Query Match 58.6%; Score 34; DB 4; Length 229;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10  
Db 200 DSLPPRDK 207

RESULT 12  
US-08-258-261B-3  
Sequence 3; Application US/08258261B  
Patent No. 5639949

GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Uknes, Scott Joseph  
TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/258,261B  
FILING DATE: 08-JUN-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,205  
FILING DATE: 01-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-258-261B-3

Query Match 58.6%; Score 34; DB 1; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPRE 9  
Db 107 NLPPRE 112

RESULT 13  
US-08-456-837-3  
Sequence 3; Application US/08456837  
Patent No. 5843774  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Uknes, Scott Joseph  
TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances

NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,837  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/457,205  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-456-837-3

Query Match 58.6%; Score 34; DB 1; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPE 9  
|||||  
DB 107 NLPPE 112

RESULT 14  
US-08-457-342-3  
Sequence 3, Application US/08457342  
Patent No. 5652898  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hamner, Phillip E.  
APPLICANT: Uknes, Scott Joseph  
TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,342  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,205  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-457-342-3

Query Match 58.6%; Score 34; DB 1; Length 361;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPE 9  
|||||  
DB 107 NLPPE 112

RESULT 15  
US-08-457-646A-3  
Sequence 3, Application US/08457646A  
Patent No. 5679560  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hamner, Phillip E.  
APPLICANT: Uknes, Scott Joseph  
TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,646A  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,205  
FILING DATE: 01-JUN-1995

APPLICATION NUMBER: 08/258,261  
 FILING DATE: 08-Jun-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 361 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-457-646A-3

Query Match 58.6% Score 34; DB 1; Length 361;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 NLPPRE 9  
 DB 107 NLPPRE 112

Search completed: June 13, 2002, 12:22:56  
 Job time: 194 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:21 ; Search time 72.45 Seconds  
(without alignments)  
16.864 Million cell updates/sec

Title: US-09-727-739B-18  
Perfect score: 58  
Sequence: 1 SVDNPPRRK 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	111	22	AAU07668
2	49	84.5	28	7	AAP61714
3	49	84.5	28	10	AAP90989
4	46	79.3	845	22	ABB2651
5	46	79.3	845	22	ABB2651
6	45	77.6	115	22	AAU07667
7	40	69.0	290	22	AAU07667
8	40	69.0	1006	22	ABB63556
9	39	67.2	219	21	AAU07849
10	39	67.2	224	21	AAU07840
11	39	67.2	244	22	AAU07843

12	39	67.2	370	22	ABBS2820	Escherichia coli p
13	39	67.2	796	22	ABB69479	Drosophila melanog
14	38	65.5	33	19	AAW79680	Synthetic CS198 de
15	38	65.5	33	22	AAE07323	Human CS 198 pepti
16	38	65.5	175	22	AAU17725	Human protein sequ
17	38	65.5	197	22	AAU17725	Novel human respir
18	38	65.5	281	21	AAU17725	Human ORFX ORF1764
19	38	65.5	486	22	ABB59435	Drosophila melanog
20	38	65.5	490	22	ABB59435	Drosophila melanog
21	37	63.8	56	22	AAU07668	Human immune haema
22	37	63.8	73	21	AAU07668	Arabidopsis thalia
23	37	63.8	76	21	AAU07668	Arabidopsis thalia
24	37	63.8	76	21	AAU07668	Arabidopsis thalia
25	37	63.8	76	21	AAU07668	Arabidopsis thalia
26	37	63.8	76	21	AAU07668	Arabidopsis thalia
27	37	63.8	76	22	AAU07668	Arabidopsis thalia
28	37	63.8	110	21	AAU07668	Arabidopsis thalia
29	37	63.8	110	21	AAU07668	Arabidopsis thalia
30	37	63.8	262	22	ABB65208	Drosophila melanog
31	37	63.8	392	22	ABB65208	Novel human diago
32	37	63.8	424	10	AAU07668	Amino acid sequenc
33	37	63.8	424	11	AAU07668	SPI-like protein e
34	37	63.8	497	22	ABB21065	Novel human diago
35	36	62.1	76	21	AAU07668	Arabidopsis thalia
36	36	62.1	228	22	AAU07668	Human polypeptide
37	36	62.1	245	22	AAU07668	Human polypeptide
38	36	62.1	268	22	AAU07668	S. epidermidis ope
39	36	62.1	287	22	ABB58433	Drosophila melanog
40	36	62.1	290	21	AAU07668	Human ORFX ORF1808
41	36	62.1	304	22	AAU07668	Human polypeptide
42	36	62.1	315	22	AAU07668	Human polypeptide
43	36	62.1	381	21	AAU07668	Drosophila melanog
44	36	62.1	381	22	ABB65532	Drosophila melanog
45	36	62.1	392	21	AAU07668	Drosophila odorant

## ALIGNMENTS

RESULT 1  
AAU07668  
ID AAU07668 standard; Protein; 111 AA.

AC AAU07668;

DT '04-DEC-2001 (first entry)

DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.

XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
XX PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
XX gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
XX carcinoma; cell proliferation; apoptosis; growth hormone;  
XX glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
XX epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
XX neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
XX anti-human immunodeficiency virus; osteopathic; anticonvulsant.

OS Oncorhynchus mykiss.

Key	Location/Qualifiers
Key	1..25
Peptide	/note- "Signal peptide"
Protein	1..86
Protein	/note- "PPSS-II' pre-sequence"
Protein	26..111
Protein	/note- "Mature PPSS-II'"
Peptide	87..97
Peptide	/note- "PPSS-II' pro-sequence"
Peptide	87..111
Peptide	/note- "Prosomatostatin II'"
Cleavage-site	96..97
Cleavage-site	/note- "Dibasic cleavage site"

Peptide 98...111  
/note= "SS-14 variant peptide"

CA2325169-A1.  
03-JUN-2001.  
01-DEC-2000; 2000CA-2325169.  
03-DEC-1999; 99US-0168934.  
(NDSU-) NDSU RES FOUND.  
Sheridan MN, Moore CA, Kittelson JD;  
WPI: 2001-425997/46.  
N-PSDB; AAS12935.

New somatostatin polypeptides derived from *Oncorhynchus mykiss*, useful for treating diabetes mellitus, acromegaly, gastrinoma, acquired immunodeficiency syndrome and neurological disorders -

Claim 1; Fig 3; 52pp; English.

The invention relates to an *Oncorhynchus mykiss* somatostatin polypeptide containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of preprosomatostatin II (PPSS-II). The protein sequences and their associated polynucleotides are useful for identifying modified somatostatin polypeptide which functions as a somatostatin agonist useful for research, therapeutics or diagnostics, including medical and veterinary applications. The wild-type somatostatin and its modified version are useful for treating hypersecretion from endocrine tumours in the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g. gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage through their effects on cell proliferation and apoptosis and as adjuvants in the treatment of diabetes mellitus via inhibition of growth hormone and glucagon. In addition, dysfunctional somatostatin secretion is associated with acquired immunodeficiency syndrome (AIDS) and various neurological disorders (e.g. epilepsy, Alzheimer's disease and Huntington's disease) and somatostatin antagonists are effective in the treatment of such conditions. Nucleic acids encoding the polypeptides are useful in gene therapy and fusion peptides can be targeted to neoplasms and their metastases, inhibiting the release of their secretory products. This sequence represents O. Mykiss PPSS-II', protein.

Note: The features for this sequence are specifically claimed in the specification.

Sequence 111 AA;

Query Match 100.0%; Score 58; DB 22; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.0046;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SVDNLPPEPRK 11  
| | | | | | | | | |  
Db 87 svdnlpprerk 97

RESULT 2  
AAP61714  
ID AAP61714 standard; Protein; 28 AA.

AC AAP61714;  
DT 28-JUL-1991 (first entry)  
XX Somatostatin-28 analogue.

XX Somatostatin-28; insulin-selective; insulinoma.  
OS Synthetic.  
XX

Key Location/Qualifiers  
Misc-difference 23 /label= Hyl, Lys  
EP173527-A.  
05-MAR-1986.  
16-AUG-1985; 85EP-0305867.  
31-AUG-1984; 84US-0646610.  
(SALK ) SALK INST FOR BIOL STUD.  
Spless J, Noe BD;  
WPI: 1986-063363/10.  
Angler fish somatostatin-28 and analogue and fragment - useful in inhibiting insulin secretion in insulinoma.  
Claim 4; Page 18; 19pp; English.  
The protein sequence is an insulin-selective analogue of anglerfish somatostatin-28, which is more potent than somatostatin-14 or somatostatin-28 in inhibiting insulin secretion for treatment of insulinoma.  
Sequence 28 AA;

Query Match 84.5%; Score 49; DB 7; Length 28;  
Best Local Similarity 81.8%; Pred. No. 0.048;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SVDNLPPEPRK 11  
| | | | | | | | | |  
Db 4 stnlppeprk 14

RESULT 3  
AAP90989  
ID AAP90989 standard; peptide; 28 AA.

AC AAP90989;  
XX  
DT 08-JUN-1990 (first entry)

XX Analogue of anglerfish somatostatin 28.

XX Somatostatin 28; SS-28; analogue; insulin secretion inhibitor; insulinoma; gastric acid secretion; thermoregulation.  
OS Anglerfish.

Key Location/Qualifiers  
Disulfide-bond 17 /note="Bonded to Cys-28"  
Disulfide-bond 28 /note="Bonded to Cys-17"  
Misc-difference 23 /label=Lys, Hyl  
Region 15..28 /note="Also claimed"

XX US4816438-A.

PN 28-MAR-1989.

XX 01-APR-1987; 87US-0033295.

XX 01-APR-1987; 87US-0033295.  
XX 31-AUG-1984; 84US-0646610.

XX PA (SALK ) SALK INST FOR BIOL STUD.  
XX PI Spiess J, Noe BD;  
XX XX WPI; 1989-113910/15.  
XX XX Angler fish somatostatin-28 and fragments -  
XX PT useful in inhibiting insulin secretion and insulinoma  
XX PS Claim 1; page 65; 8pp; English.  
XX CC It is called ASS-28 because it is an analogue of anglerfish somatostatin  
XX CC (SS-28). It is more potent than either somatostatin 14 (SS-14) or SS-28  
XX CC at inhibiting insulin secretion for the treatment of insulinoma. The  
XX CC 14-residue C-terminal peptide is also claimed (ASS-14). ASS-14 is useful  
XX CC for inhibiting insulin secretion by the pancreas. ASS-28 and ASS-28 may  
XX CC be useful for decreasing gastric acid secretion and influencing  
XX CC thermoregulation. Their reduced linear forms, wherein the disulphide  
XX CC bridge is not present and is replaced by H, is also claimed.  
XX SQ Sequence 28 AA;

Query Match 84.5%; Score 49; DB 10; Length 28;  
Best Local Similarity 81.8%; Pred. No. 0.048;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11  
DB 4 stnnlpprerk 14

RESULT 4  
AAP20029  
ID AAP20029 standard; Protein; 125 AA.  
XX AC AAP20029;  
XX DT 14-AUG-1992 (first entry)  
XX XX Sequence of preprosomatostatin-2 encoded on plas2.  
XX DE Somatostatin; growth hormone; peptide hormone; secretion.  
XX KW Somatostatin; growth hormone; peptide hormone; secretion.  
XX FH Key Location/Qualifiers  
XX FT Protein 112..125  
XX T /label= Somatostatin II  
XX EP46669-A.  
XX PN  
XX PD 03-MAR-1982.  
XX XX  
XX PF 21-AUG-1981; 81EP-0303825.  
XX PR 25-AUG-1980; 80US-0181046.  
XX XX  
XX PA (REGC ) UNIV OF CALIFORNIA;  
XX XX Hobart P, Crawford R, Pictet RL, Rutter WJ;  
XX PI WPI; 1982-18113E/10.  
XX DR N-PSDB; AAN20034.  
XX XX  
XX PT New somatostatin and precursors - produced by transformed  
XX PT microorganisms  
XX PS Example; Fig 3; 50pp; English.  
XX XX  
XX CC The inventors claim preprosomatostatin-1, prosomatostatin-1,  
XX CC preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA  
XX CC encoding them. The translation of somatostatin mRNA yields a  
XX CC precursor (prepro S1) containing a signal peptide which may be

CC released during the transit into the endoplasmic reticulum, and the  
CC resultant precursor (pro S1) is subsequently cleaved to yield S1  
CC itself. The prepeptide portion of prepro S1 is probably about 20-25  
CC bases long. Translation of plas2 predicts the sequence of a 125 AA  
CC peptide which surprisingly contains a 14 AA sequence at its carboxy  
CC terminus which differs from S1 by only 2 AAs, and is termed  
CC Somatostatin 2 (S2).  
XX SQ Sequence 125 AA;

Query Match 84.5%; Score 49; DB 3; Length 125;  
Best Local Similarity 81.8%; Pred. No. 0.23;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEPRK 11  
DB 101 stnnlpprerk 111

RESULT 5  
ABB62651  
ID ABB62651 standard; Protein; 845 AA.  
XX AC ABB62651;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 14745.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmaceutical.  
XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US09231.  
XX PR 23-MAR-2000; 2000US-191637P.  
XX PR 11-JUL-2000; 2000US-0614150.  
XX PA (PEKE ) PE CORP NY.  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI; 2001-656860/75.  
XX DR N-PSDB; ABL06734.  
XX XX  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions -  
XX PS Disclosure; SEQ ID NO 14745; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX CC sequences (AB101840-AB16175) and the encoded proteins  
XX CC (ABE57737-ABE72072).  
XX CC The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 845 AA;

Query Match 79.3%; Score 46; DB 22; Length 845;



PT Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -  
 XX Example 3; SEQ ID NO 2085; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AA138642-AA142213) with neurotropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 XX  
 SQ Sequence 290 AA;

Query Match 69.0%; Score 40; DB 22; Length 290;  
 Best Local Similarity 72.7%; Pred. No. 24;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEK 11  
 DB 193 swdsippserk 203  
 I :||| |||

RESULT 8  
 ID ABB63556 standard; Protein; 1006 AA.  
 XX  
 AC ABB63556;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 17460.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 JS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 DR WPI; 2001-65860/75.  
 DR N-PSDB; ABL07659.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 17460; 21pp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 1006 AA;

Query Match 69.0%; Score 40; DB 22; Length 1006;  
 Best Local Similarity 60.0%; Pred. No. 87;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPEK 10  
 DB 444 atdmpprkr 453  
 : ||||| :

RESULT 9  
 ID AAB07849 standard; Protein; 219 AA.  
 XX  
 AC AAB07849;  
 XX  
 DT 14-NOV-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soybean type III glutathione-S-transferase.

XX Soybean; glutathione-S-transferase; GST; detoxification;  
 KW xenobiotic compound; herbicide-tolerance; transgenic plant;  
 KW herbicide synergist.  
 XX  
 OS Glycine max.  
 XX  
 PN WO2000047728-A2.  
 XX  
 PD 17-AUG-2000.  
 XX  
 PF 10-FEB-2000; 2000WO-US03347.  
 XX  
 PR 10-FEB-1999; 99US-0247373.  
 XX  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI McGonigle B, O'Keefe DP;  
 XX  
 DR WPI; 2000-549144/50.  
 DR N-PSDB; AAA59484.  
 XX  
 XX Soybean glutathione-S-transferase polypeptides and polynucleotides used  
 PT to produce herbicide tolerant transgenic plants and to screen for  
 PT inhibitors or substrates of the enzyme -  
 XX  
 PS Claim 26; Page 81-82; 84pp; English.

XX The present sequence represents a soybean glutathione-S-transferase  
 CC (GST) enzyme. The enzyme is involved in the detoxification of  
 CC xenobiotic compounds in plants and seeds. The GST polynucleotides  
 CC and polypeptides are used for the production of herbicide-tolerant  
 CC transgenic plants, and for the development of screening assays to  
 CC identify GST inhibitors and substrates, which can be used as  
 CC herbicide synergists. GST Gene specific probes can be used in gene  
 CC identification methods. The recombinant GST enzymes can be used to  
 CC produce enzyme specific antibodies which are used to detect the  
 CC enzymes in situ in cells or in vitro in cell extracts.  
 XX  
 XX Sequence 219 AA;



Query Match 57.2%; Score 39; DB 21; Length 219;  
 Best Local Similarity 75.0%; Pred. No. 27;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10  
 Db 194 dnlpprdk 201  
 |||||:

RESULT 10  
 AAB07840  
 ID AAB07840 standard; Protein; 224 AA.  
 XX  
 AC AAB07840;  
 XX  
 DT 14-NOV-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soybean type III glutathione-S-transferase.  
 XX  
 KW Soybean; glutathione-S-transferase; GST; detoxification;  
 KW xenobiotic compound; herbicide-tolerance; transgenic plant;  
 W herbicide synergist.  
 XX  
 OS Glycine max.  
 XX  
 PN WO200047728-A2.  
 XX  
 PD 17-AUG-2000.  
 XX  
 PF 10-FEB-2000; 2000WO-US03347.  
 XX  
 PR 10-FEB-1999; 99US-0247373.  
 XX  
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.  
 XX  
 PI McGonigle B, O'Keefe DP;  
 XX  
 DR WPI; 2000-549144/50.  
 DR N-PSDB; AAA59475.  
 XX  
 XX Soybean glutathione-S-transferase polypeptides and polynucleotides used  
 PT to produce herbicide tolerant transgenic plants and to screen for  
 PT inhibitors or substrates of the enzyme -  
 PS Claim 26; Page 71-72; 84pp; English.  
 XX  
 CC The present sequence represents a soybean glutathione-S-transferase  
 CC (GST) enzyme. The enzyme is involved in the detoxification of  
 CC xenobiotic compounds in plants and seeds. The GST polynucleotides  
 CC and polypeptides are used for the production of herbicide-tolerant  
 CC transgenic plants and for the development of screening assays to  
 CC identify GST inhibitors and substrates, which can be used as  
 CC herbicide synergists. GST gene specific probes can be used in gene  
 CC identification methods. The recombinant GST enzymes can be used to  
 CC produce enzyme specific antibodies which are used to detect the  
 CC enzymes in situ in cells or in vitro in cell extracts.  
 XX  
 SQ Sequence 224 AA;

Query Match 67.2%; Score 39; DB 21; Length 224;  
 Best Local Similarity 75.0%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10  
 Db 202 dnlpprdk 209  
 |||||:

RESULT 11  
 AAG82843  
 ID AAG82843 standard; Protein; 244 AA.  
 XX

AC AAG82843;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX  
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2780.  
 XX  
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis;  
 KW vaccination; endocarditis.  
 XX  
 OS Staphylococcus epidermidis.  
 XX  
 PN WO200134809-A2.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 09-NOV-2000; 2000WO-US30782.  
 XX  
 PR 09-NOV-1999; 99US-0164258.  
 XX  
 PA (GLAX) GLAXO GROUP LTD.  
 XX  
 PI Kimmerly WJ;  
 XX  
 DR WPI; 2001-316495/33.  
 DR N-PSDB; AAH53693.  
 XX  
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
 PT useful for vaccinating against infections, e.g. endocarditis -  
 XX  
 PS Claim 18; Page 724; 2188pp; English.  
 XX  
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAH51454 to AAH53120, from Staphylococcus epidermidis.  
 CC (I) and (II) can have antibacterial activity and therefore can be used  
 CC in vaccination. The nucleic acids (I) may be used to produce the  
 CC S. epidermidis polypeptides (II) via the production of vectors  
 CC containing them which are used to produce host cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 CC used to vaccinate subjects and to raise antibodies against the bacteria.  
 CC The polypeptides may also be used to assay for other inhibitors of their  
 CC activity and therefore identify compounds that may be used for the  
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA  
 CC polynucleotide sequences from the present invention. AAH55091 to  
 CC AAH55098 represent oligonucleotide sequences and primers which are used  
 CC in the exemplification of the present invention.  
 CC N.B. The present invention specifically claims all the polynucleotide  
 CC sequences given in the sequence listing of the present specification,  
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,  
 CC no sequences are present for SEQ ID NO:4455 to 4464.  
 XX  
 SQ Sequence 244 AA;

Query Match 67.2%; Score 39; DB 22; Length 244;  
 Best Local Similarity 70.0%; Pred. No. 30;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPIR 10  
 Db 177 saknlppkr 186  
 |||||:

RESULT 12  
 ABB52820  
 ID ABB52820 standard; Protein; 370 AA.  
 XX  
 AC ABB52820;  
 XX  
 DT 11-FEB-2002 (first entry)  
 XX  
 DE Escherichia coli polypeptide SEQ ID NO 1044.

XX Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;  
 KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;  
 KW systemic infection; non-diarrhoeal infection; septicaemia;  
 KW pyelonephritis; antibiotic resistance.

XX Escherichia coli.

PN WO200166572-A2.

XX 13-SEP-2001.

XX 12-MAR-2001; 2001WO-EP03445.

XX 10-MAR-2000; 2000FR-0003145.

PR 02-FEB-2001; 2001FR-0001449.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;

XX WPI; 2001-550253/61.

XX A library of DNA fragments of Escherichia coli strains for the  
 PT phylogenetic determination of a given strain comprises polynucleotides of  
 PT nature B2/D+ A- -

XX Example 6; Fig 6; 646pp; English.

XX The invention relates to a library of DNA fragments of Escherichia coli  
 CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)  
 CC and encoded proteins (AB52459-AB52919 and AB52954-AB53094) of nature  
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,  
 CC antibacterial and immunosuppressive activity as part of pharmaceutical  
 CC compositions used to treat, palliate or prevent extra-intestinal E. coli  
 CC infections. The polypeptides are useful for determining the phylogenetic  
 CC group of a given E. coli strain. These polypeptides can detect and treat  
 CC an undesired development of E. coli, particularly an extra-intestinal  
 CC infection that include systemic and non-diarrhoeal infections such as  
 CC septicaemia, pyelonephritis and meningitis this is particularly  
 CC advantageous as bacterial resistance is increasing with the more  
 CC frequent use of broad spectrum antibiotics.

XX Sequence 370 AA;

Query Match 67.2%; Score 39; DB 22; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDNLPFR 8  
 DB 169 vdnlpfr 175

RESULT 13

ABB69479  
 ID ABB69479 standard; Protein; 796 AA.

XX ABB69479;

XX 26-MAR-2002 (first entry).

XX Drosophila melanogaster polypeptide SEQ ID NO 35229.

XX Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL13582.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -

XX Disclosure; SEQ ID NO 35229; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 796 AA;

Query Match 67.2%; Score 39; DB 22; Length 796;  
 Best Local Similarity 66.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRRK 11  
 DB 786 nmlpprrr 794

RESULT 14

AAW79680

ID AAW79680 standard; Protein; 33 AA.

XX AAW79680;

XX 11-JAN-1999 (first entry)

XX Synthetic CSI98 derived peptide #4.

XX Gastrointestinal tract; GI tract; cancer; disease; detection; CSI98;  
 KW human; predisposition; treatment; Barrett's oesophagus; gastric ulcer;  
 KW gastritis; leiomyoma; polyps; Crohn's disease; ulcerative colitis;  
 KW pancreatitis.

XX Synthetic.

XX WO9844159-A1.

XX 08-OCT-1998.

XX 30-MAR-1998; 98WO-US06251.

XX 31-MAR-1997; 97US-0828855.

XX (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;

PI Gordon J, Granados EN, Hayden M, Hodges SC, Klass MR;

PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD;

XX DR WPI; 1998-542714/46.  
 XX PT New gastrointestinal polynucleotides, CS198, and their detection -  
 PT used for developing products for the diagnosis and treatment of  
 PT gastrointestinal disorders, e.g. cancers, gastric ulcer or gastritis  
 XX PS  
 XX Claim 26; Page 103; 127pp; English.  
 XX  
 XX AAW79677-W79680 are synthetic CS198 derived peptide fragments which are  
 CC used in a method to detect the presence of a target human CS198  
 CC polynucleotide in a test sample. The CS198 gene is useful as a marker  
 CC for gastrointestinal (GI) tract disorders. The methods and products can  
 CC be used in detecting, diagnosing, staging, monitoring, prognosticating,  
 CC preventing or treating, or determining the predisposition to diseases  
 CC and conditions of the GI tract, such as GI tract cancer, Barrett's  
 CC esophagus, gastric ulcer, gastritis, leiomyoma, polyps, Crohn's disease,  
 CC ulcerative colitis, and pancreatitis.  
 XX SQ Sequence 33 AA;

Query Match 65.5%; Score 38; DB 19; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 5.7;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SVDNLPPEPRK 11  
 |||: ||: ||  
 Db 3 svdsappgqrk 13

## RESULT 15

AAE07323  
 ID AAE07323 standard; peptide; 33 AA.  
 XX  
 XX AAE07323;  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human CS 198 peptide #4.  
 XX  
 XX CS 198; gastrointestinal tract disease; GI tract; cancer; gastric ulcer;  
 KW gastritis; Crohn's disease; ulcerative colitis; pancreatitis;  
 KW Barrett's esophagus; gene therapy; drug screening; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2001010904-A1.  
 XX  
 XX 02-AUG-2001.  
 XX  
 XX 30-MAR-1998; 98US-0050516.  
 XX  
 XX 31-MAR-1997; 97US-0828855.  
 XX  
 XX (BILL/) BILLING-MEDEL P A.  
 XX (COHE/) COHEN M.  
 XX (COLP/) COLPITTS T L.  
 XX (FRIE/) FRIEDMAN P N.  
 XX (GORD/) GORDON J.  
 XX (GRAN/) GRANADOS E N.  
 XX (HAYD/) HAYDEN M.  
 XX (HODG/) HODGES S C.  
 XX (KLAS/) KLAS M R.  
 XX (KRAT/) KRATOCHVIL J D.  
 XX (ROBE/) ROBERTS-RAPP L.  
 XX (RUSS/) RUSSELL J C.  
 XX (STRO/) STROUPE S D.  
 XX  
 XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;  
 PI Granados EN, Hayden M, Hodges SC, KLAS MR, Kratochvil JD;  
 PI Roberts-Rapp L, Russell JC, Stroupe SD;  
 XX

DR WPI; 2001-496163/54.

XX  
 XX Detecting the presence of target CS 198 polynucleotide, useful for  
 PT detecting or diagnosing diseases of the gastrointestinal tract,  
 PT comprises contacting test sample with at least one CS 198-specific  
 PT polynucleotide  
 XX  
 XX Claim 17; Page 52; 68pp; English.

XX  
 XX The invention relates to a method of detecting the presence of a target  
 CC CS 198 polynucleotide comprising contacting the test sample with at  
 CC least one CS 198-specific polynucleotide. The method is useful for  
 CC detecting diseases of the gastrointestinal (GI) tract organs,  
 CC particularly cancer. The CS 198 polynucleotides, polypeptides and  
 CC antibodies are useful for detecting, diagnosing, staging, monitoring,  
 CC prognosticating, preventing, treating or determining predisposition to  
 CC diseases and conditions of the GI tract such as cancer, gastric ulcer,  
 CC gastritis, Crohn's disease, ulcerative colitis, pancreatitis and  
 CC Barrett's esophagus. The CS 198 polypeptides are useful as standards  
 CC or reagents in diagnostic immunoassays, as components or as  
 CC target sites for various therapies. Antibodies directed against at  
 CC least one epitope contained within these polypeptides are useful as  
 CC delivery agents for therapeutic agents, in diagnostic tests and for  
 CC screening for conditions or diseases associated with CS 198,  
 CC particularly cancer. Monoclonal antibodies may also be used for the  
 CC generation of chimeric antibodies for therapeutic use. The CS 198  
 CC polynucleotide is also useful in gene therapy and drug screening.  
 CC The method of the invention provides an alternative, non-surgical  
 CC diagnostic method capable of detecting early stage GI tract disease  
 CC such as cancer. The present sequence is a peptide derived from  
 CC human CS 198 polypeptide consensus sequence.  
 XX SQ Sequence 33 AA;

Query Match 65.5%; Score 38; DB 22; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 5.7;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SVDNLPPEPRK 11  
 |||: ||: ||  
 Db 3 svdsappgqrk 13

Search completed: June 13, 2002, 12:22:22  
 Job time: 275 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:17:47 ; Search time 72.45 Seconds  
(without alignments)  
174.775 Million cell updates/sec

Title: US-09-727-739B-3  
Perfect score: 566  
Sequence: 1 MLSTRVOCALSLSLAIAIS.....APRRKAGCKNFNFKTFTSC.114

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566	100.0	114	22 AAU07666	Rainbow trout prep
2	360	63.6	116	21 AAG03774	Human secreted pro
3	251	44.3	121	3 AAP20028	Sequence of prepro
4	177.5	31.4	115	22 AAU07667	Rainbow trout prep
5	164.5	29.1	111	22 AAU07668	Rainbow trout prep
6	162.5	28.7	125	3 AAP20029	Sequence of prepro
7	127	22.4	25	3 AAP20198	Sequence of somato
8	127	22.4	25	22 AAB91017	Somatostatin relat
9	127	22.4	28	19 AAW51859	Somatostatin analo
10	127	22.4	28	20 AAY28703	Mouse somatostatin
11	127	22.4	28	20 AAY24384	Somatostatin pepti

12	127	22.4	28	20 AAY24339	Peptide hormone so
13	127	22.4	28	22 AAB91018	Somatostatin relat
14	127	22.4	28	22 AAU07669	Mammalian somatost
15	127	22.4	29	22 AAB91020	Somatostatin relat
16	127	22.4	140	21 AAY69789	MWpSp-MWpMp20-(His
17	126	22.3	28	19 AAW50950	Somatostatin analo
18	126	22.3	28	22 AAB91021	Somatostatin relat
19	123	21.7	28	19 AAW50816	Somatostatin-28 an
20	122	21.6	28	3 AAP20130	Somatostatin-28 pe
21	122	21.6	28	3 AAP20131	Somatostatin-28 pe
22	120	21.2	28	3 AAP20137	Sequence of [D-Trp
23	119	21.0	28	19 AAW50818	Somatostatin-28 an
24	119	21.0	28	19 AAW50819	Somatostatin-28 an
25	119	21.0	28	19 AAW51855	Somatostatin analo
26	118	20.8	20	6 AAP50414	Swine duodenum eic
27	118	20.8	27	22 AAB91007	Somatostatin relat
28	118	20.8	28	4 AAP30058	Radioactively tagg
29	117	20.7	28	3 AAP20125	Somatostatin-28 pe
30	117	20.7	28	19 AAW51688	Somatostatin analo
31	113	20.0	28	3 AAP20126	Somatostatin-28 pe
32	111	19.6	28	3 AAP20128	Somatostatin-28 an
33	111	19.6	28	19 AAW50817	Somatostatin-28 an
34	101	17.8	28	7 AAP61714	Analogue of angler
35	101	17.8	28	10 AAP90989	Somatostatin relat
36	101	17.8	32	22 AAB91009	Somatostatin-28 pe
37	100	17.7	28	3 AAP20127	Somatostatin-28 pe
38	95	16.8	943	14 AAR34346	Somatostatin-leuko
39	94.5	16.7	112	19 AAW42033	Rat preprocorcista
40	93.5	16.5	175	10 AAP90494	Dihydrofolate redu
41	93.5	16.5	175	10 AAP90496	Dihydrofolate redu
42	93	16.4	36	21 AAY58743	Somatostatin antig
43	89.5	15.8	36	21 AAY58760	Somatostatin antig
44	89.5	15.8	36	21 AAY58751	Somatostatin antig
45	89	15.7	24	20 AAY39551	CTLA-4 VLD CDR1 re

ALIGNMENTS

RESULT 1  
AAU07666  
ID AAU07666 standard; Protein: 114 AA.  
XX AC AAU07666;  
XX DT 04-DEC-2001 (first entry)  
XX DE Rainbow trout preprosomatostatin I (PPSS-I) polypeptide.

Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
anti-human immunodeficiency virus; osteopathic; anticonvulsant.

Key	Location/Qualifiers
Peptide	/note= "Signal peptide"
Protein	/note= "PPSS-I pre-sequence"
Protein	25..114
Protein	/note= "Mature PPSS-I"
Peptide	89..100
Peptide	/note= "PPSS-I pro-sequence"
Peptide	89..114
Peptide	/note= "Prosomatostatin I"
Cleavage-site	99..100
	/note= "Dibasic cleavage site"

FT Peptide 101..114  
XX /note= "SS-14 peptide"  
PN CA2325169-A1.  
XX 03-JUN-2001.  
PD 01-DEC-2000; 2000CA-2325169.  
XX 03-DEC-1999; 99US-0168934.  
XX (NDSU-) NDSU RES FOUND.  
XX Sheridan MA, Moore CA, Kittelson JD;  
XX WPI; 2001-425997/46.  
XX N-PSDB; RAS12933.  
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
XX immunodeficiency syndrome and neurological disorders -  
XX Claim 1; Fig 2; 52pp; English.  
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
XX preprosomatostatin II (PPSS-II). The protein sequences and their  
XX associated polynucleotides are useful for identifying modified  
XX somatostatin polypeptides which functions as a somatostatin agonist useful  
XX for research, therapeutics or diagnostics, including medical and  
XX veterinary applications. The wild-type somatostatin and its modified  
XX version are useful for treating hypersecretion from endocrine tumours in  
XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
XX through their effects on cell proliferation and apoptosis and as adjuncts  
XX in the treatment of diabetes mellitus via inhibition of growth hormone  
XX and glucagon. In addition, dysfunctional somatostatin secretion is  
XX associated with acquired immunodeficiency syndrome (AIDS) and various  
XX neurological disorders (e.g. epilepsy, Alzheimer's disease and  
XX Huntington's disease) and somatostatin antagonists are effective in the  
XX treatment of such conditions. Nucleic acids encoding the polypeptides are  
XX useful in gene therapy and fusion peptides can be targeted to neoplasms  
XX and their metastases, inhibiting the release of their secretory products.  
XX This sequence represents O. Mykiss PPSS-I protein.  
XX Note: The features for this sequence are specifically claimed in the  
XX specification.  
XX Sequence 114 AA;  
Query Match 100.0%; Score 566; DB 22; Length 114;  
Best Local Similarity 100.0%; Pred. No. 2.8e-55;  
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLSTRVOCALALLSLALATSSVSAAPSDAKLRQLLRSLMAPAGKQELARNTLVLLSEL 60  
DB 1 mlstrvocalallslalatsvsaapsdaklrqlrrslmapagkgelartlvllsel 60  
QY 61 AHVNEAIEIDMSHGVEQEDVDLELERAPGVLPAPRERKAGCKNFFWKTFTSC 114  
DB 61 ahvneaieidmsghveqedvdlelerapgvlpaprerkgacknffwktftsc 114  
RESULT 2  
AAG03774  
ID AAG03774 standard; Protein; 116 AA.  
XX AAG03774;  
XX 06-OCT-2000 (first entry)  
XX Human secreted protein, SEQ ID NO: 7855.  
XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.  
XX Homo sapiens.  
XX EP1033401-A2.  
XX 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-0200610.  
XX 26-FEB-1999; 99US-0122487.  
XX (GEST ) GENSET.  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX N-PSDB; AAC03780.  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX Claim 13; SEQ ID 7855; 71pp + CD-ROM; English.  
XX The present sequence is a polypeptide encoded by one of a large number  
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
XX were prepared from total human RNAs or polyA+ RNAs derived from 30  
XX different tissues. EST sequences usually correspond mainly to the 3'  
XX untranslated region (UTR) of the mRNA because they are often obtained  
XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
XX isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
XX those cases where longer cDNA sequences have been obtained, the full 5'  
XX UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
XX ends and can therefore be used to obtain full length cDNAs and genomic  
XX DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
XX chromosome mapping procedures. They are used to obtain upstream  
XX regulatory sequences and to design expression and secretion vectors.  
XX Sequence 116 AA;  
Query Match 63.6%; Score 360; DB 21; Length 116;  
Best Local Similarity 62.1%; Pred. No. 2.5e-32;  
Matches 72; Conservative 17; Mismatches 25; Indels 2; Gaps 1;  
QY 1 MLSTRVOCALALLSLALATSSVSAAPSDAKLRQLLRSLMAPAGKQELARNTLVLLSEL 60  
DB 1 mlstrvocalallslalatsvsaapsdaklrqlrrslmapagkgelartlvllsel 60  
QY 61 AHVNEAIEIDMSHGVEQEDVDLELERAPGVLPAPRERKAGCKNFFWKTFTSC 114  
DB 61 ahvneaieidmsghveqedvdlelerapgvlpaprerkgacknffwktftsc 116  
RESULT 3  
AAP20028  
ID AAP20028 standard; Protein; 121 AA.  
XX AAP20028;  
XX 14-AUG-1992 (first entry)  
XX Sequence of preprosomatostatin-1 encoded on plasmid.  
XX Somatostatin; growth hormone; peptide hormone; secretion.  
XX Key Location/Qualifiers  
XX Protein 108..121  
XX /label= Somatostatin I  
XX EP46669-A.

```

XX PD 03-MAR-1982.
XX PF 21-AUG-1981; 81EP-0303825.
XX PR 25-AUG-1980; 80US-0181046.
XX PA (REGC ) UNIV OF CALIFORNIA.
XX PI Hobart P, Crawford R, Pictet RL, Rutter WJ;
XX WPI: 1982-18113E/10.
XX N-PSDB; AAN20033.
XX New somatostatin and precursors - produced by transformed
XX microorganisms
XX Example: Fig 3; 50pp; English..
XX The inventors claim preprosomatostatin-1, prosomatostatin-1,
XX preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA
XX encoding them. The translation of somatostatin mRNA yields a
XX precursor (prepro S1) containing a signal peptide which may be
XX released during the transit into the endoplasmic reticulum, and the
XX resultant precursor (pro S1) is subsequently cleaved to yield S1
XX itself. The prepeptide portion of prepro S1 is probably about 20-25
XX bases long. Translation of p1a52 predicts the sequence of a 125 AA
XX peptide which surprisingly contains a 14 AA sequence at its carboxy
XX terminus which differs from S1 by only 2 AAs, and is termed
XX Somatostatin 2 (S2).
XX Sequence 121 AA;
XX
XX Query Match 44.3%; Score 251; DB 3; Length 121;
XX Best Local Similarity 51.3%; Pred. No. 3.6e-20;
XX Matches 60; Conservative 15; Mismatches 36; Indels 6; Gaps 4;
XX
XX QY 3 STRVOCALA-LLSLALAISSVSAAPSADKLRLQLRSLMAPAGKQELARNTLVE-LISEL 60
XX DB 6 ssvrlclvllslstasiscsfaggrdkirlllhyrpl-qgskqgdmtrsalaelllsdl 64
XX QY 61 AHVENEAIELDMW---SHGVQEDVDLELERAPGVLAPRRKAGCKNFFKFTFSC 114
XX DB 65 lqgenealeenfp1aeggpedahadleraaaggp1laprekrkagcknffkftsc 121
XX
XX --RESULT 4
XX AU07667
XX AU07667 standard; Protein; 115 AA.
XX AC AU07667;
XX DX 04-DEC-2001 (first entry)
XX DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
XX
XX KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
XX KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
XX KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
XX KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
XX KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
XX KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
XX KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
XX KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
XX OS Oncorhynchus mykiss.
XX
XX Key Location/Qualifiers
XX Peptide 1..25 "Signal peptide"
XX FT 1..87 "PPSS-II' pre-sequence"
XX Protein

```

```

FT Protein 26..115
FT /note= "Mature PPSS-II'"
FT Misc-difference 74
FT /note= "Encoded by CAA"
FT Peptide 88..101
FT /note= "PPSS-II' pro-sequence"
FT Peptide 88..115
FT /note= "Prosomatostatin II'"
FT Cleavage-site 100..101
FT /note= "Dibasic cleavage site"
FT Peptide 102..115
FT /note= "SS-14 variant peptide"
XX CA2325169-A1.
XX
XX 03-JUN-2001.
XX
XX 01-DEC-2000; 2000CA-2325169.
XX
XX 03-DEC-1999; 99US-0168934.
XX (NDSU-) NDSU RES FOUND.
XX Sheridan MA, Moore CA, Kittelson JD;
XX WPI: 2001-425997/46.
XX N-PSDB; AAS12934.
XX
XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired
XX immunodeficiency syndrome and neurological disorders -
XX
XX Claim 2; Fig 3; 52pp; English.
XX
XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
XX preprosomatostatin II (PPSS-II). The protein sequences and their
XX associated polynucleotides are useful for identifying modified
XX somatostatin polypeptides which functions as a somatostatin agonist useful
XX for research, therapeutics or diagnostics, including medical and
XX veterinary applications. The wild-type somatostatin and its modified
XX version are useful for treating hypersecretion from endocrine tumours in
XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
XX through their effects on cell proliferation and apoptosis and as adjuncts
XX in the treatment of diabetes mellitus via inhibition of growth hormone
XX and glucagon. In addition, dysfunctional somatostatin secretion is
XX associated with acquired immunodeficiency syndrome (AIDS) and various
XX neurological disorders (e.g. epilepsy, Alzheimer's disease and
XX Huntington's disease) and somatostatin antagonists are effective in the
XX treatment of such conditions. Nucleic acids encoding the polypeptides are
XX useful in gene therapy and fusion peptides can be targeted to neoplasms
XX and their metastases, inhibiting the release of their secretory products.
XX This sequence represents O. mykiss PPSS-II' protein.
XX Note: The features for this sequence are specifically claimed in the
XX specification.
XX
XX Sequence 115 AA:
XX
XX Query Match 31.4%; Score 177.5; DB 22; Length 115;
XX Best Local Similarity 34.7%; Pred. No. 5.2e-12;
XX Matches 50; Conservative 9; Mismatches 18; Indels 67; Gaps 4;
XX
XX QY 5 RVOCALALSLALAISSVSAAPS---DAKLRLQLRSLMA----- 41
XX DB 5 rihcalalglalalcsggaasqpdlrlsrlilqlaraaaalphrsgvserwrtfypncp 64
XX QY 42 -----PAGKQELARNTLVELLSLAHVEAIELDDMSHGVSQEDVDLELERAPCP 92
XX DB 65 clprkvpkpag----- 114
XX QY 93 V--LAPRERKAGCKNFFKFTFSC 114

```

Db 92 pnnlprckagcknfykftsc 115  
 RESULT 5  
 AAU07668 standard; Protein; 111 AA.  
 AC AAU07668;  
 DT 04-DEC-2001 (first entry)  
 DE Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.  
 KW Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
 KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
 KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
 KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;  
 KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
 KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
 KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;  
 KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.  
 OS Oncorhynchus mykiss.  
 PH Key Location/Qualifiers  
 FT Peptide 1..25 /note= "Signal peptide"  
 FT Protein 1..86 /note= "PPSS-II' pre-sequence"  
 FT Protein 26..111 /note= "Mature PPSS-II'"  
 FT Peptide 87..97 /note= "PPSS-II' pro-sequence"  
 FT Peptide 87..111 /note= "Prosomatostatin II'"  
 FT Cleavage-site 96..97 /note= "Basic cleavage site"  
 FT Peptide 98..111 /note= "SS-14 variant peptide"  
 CA2325169-A1.  
 03-JUN-2001.  
 01-DEC-2000; 2000CA-2325169.  
 03-DEC-1999; 99US-0168934.  
 (NDSU-) NDSU RES FOUND.  
 A Sheridan WA, Moore CA, Kittelson JD;  
 PI WPI: 2001-425997/46.  
 DR N-PSDB; AAS12935.  
 XX New somatostatin polypeptides derived from *Oncorhynchus mykiss*, useful  
 PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
 PT immunodeficiency syndrome and neurological disorders -  
 PS Claim 1; Fig 3; 52pp; English.  
 CC The invention relates to an *Oncorhynchus mykiss* somatostatin polypeptide  
 CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
 CC preprosomatostatin II (PPSS-II). The protein sequences and their  
 CC associated polynucleotides are useful for identifying modified  
 CC somatostatin polypeptide which functions as a somatostatin agonist useful  
 CC for research, therapeutics or diagnostics, including medical and  
 CC veterinary applications. The wild-type somatostatin and its modified  
 CC version are useful for treating hypersecretion from endocrine tumours in  
 CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
 CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
 CC through their effects on cell proliferation and apoptosis and as adjuncts

CC in the treatment of diabetes mellitus via inhibition of growth hormone  
 CC and glucagon. In addition, dysfunctional somatostatin secretion is  
 CC associated with acquired immunodeficiency syndrome (AIDS) and various  
 CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
 CC Huntington's disease) and somatostatin antagonists are effective in the  
 CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
 CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
 CC and their metastases, inhibiting the release of their secretory products.  
 CC This sequence represents O. Mykiss PPSS-II' protein.  
 CC Note: The features for this sequence are specifically claimed in the  
 CC specification.  
 XX Sequence 111 AA;  
 SQ  
 Query Match 29.1%; Score 164.5; DB 22; Length 111;  
 Best Local Similarity 38.7%; Pred. No. 1.4e-10;  
 Matches 48; Conservative 15; Mismatches 38; Indels 23; Gaps 5;  
 QY 1 MLSTRVQCALALLSLALAISSVSAAPS---DAKRLQRLQSLMAP---AGKQELARNTL 53  
 Db 1 mrvsqhcalallglalaicsgaaspqdlasrrllqralaaalphrsgvserwrtfy 60  
 QY 54 VE---LLSELAHVENEATEILDMSHGVEQEDVDLELERAGPVLAPRERKAGCKNFTWKT 110  
 Db 61 pncpcrlwrprkvkqgpk-----akeditersvdnlp-----prerkagcknfywkg 107  
 QY 111 FTSC 114  
 Db 108 ftsc 111  
 RESULT 6  
 AAP20029  
 ID AAP20029 standard; Protein; 125 AA.  
 XX  
 AC AAP20029;  
 DT 14-AUG-1992 (first entry)  
 DE Sequence of preprosomatostatin-2 encoded on plas2.  
 KW Somatostatin; growth hormone; peptide hormone; secretion.  
 PH Key Location/Qualifiers  
 FT Protein 112..125 /label= Somatostatin II  
 XX EP46669-A.  
 XX 03-MAR-1982.  
 XX 21-AUG-1981; 81EP-0303825.  
 XX 25-AUG-1980; 80US-0181046.  
 XX (REGC ) UNIV OF CALIFORNIA.  
 XX Hobart P, Crawford R, Pictet RL, Rutter WJ;  
 XX WPI: 1982-18113E/10.  
 XX N-PSDB; AAN20034.  
 XX New somatostatin and precursors - produced by transformed  
 PT microorganisms  
 PS Example; Fig 3; 50pp; English.  
 CC The inventors claim preprosomatostatin-1, preprosomatostatin-1,  
 CC preprosomatostatin-2, preprosomatostatin-2 and somatostatin-2; and DNA  
 CC encoding them. The translation of somatostatin mRNA yields a  
 CC precursor (prepro SI) containing a signal peptide which may be  
 CC released during the transit into the endoplasmic reticulum, and the



CC resultant precursor (pro S1) is subsequently cleaved to yield S1  
 CC itself. The prepeptide portion of prepro S1 is probably about 20-25  
 CC bases long. Translation of plas2 predicts the sequence of a 125 AA  
 CC peptide which surprisingly contains a 14 AA sequence at its carboxy  
 CC terminus which differs from S1 by only 2 AAs, and is termed  
 CC Somatostatin 2 (S2).

XX SQ Sequence 125 AA;  
 Query Match 28.7%; Score 162.5; DB 3; Length 125;  
 Best Local Similarity 38.9%; Pred. No. 2.7e-10;  
 Matches 51; Conservative 16; Mismatches 41; Indels 23; Gaps 6;  
 Qy 1 MLSTRVQCALALLSLAIAISSVA-----APSDAKLRQ--LLQSLMAPACKQELA 49  
 Db 1 mqdrcpallallaivcgvsvsqldregsdnqldlelrqhwllersagllsgews 60  
 Qy 50 RNTLVLLSEL----AHVNERAIELDDMSHGVEQEDVDLELERAGPV--LAPRRKAGC 103  
 b 61 kraveellagmelpatfgrea---edasmategr---mnlersvdstnlprrerkagc 114  
 Qy 104 KNFFWKTFTSC 114  
 Db 115 knfywkgftsc 125

RESULT 7  
 AAP20198  
 ID AAP20198 standard; Protein; 25 AA.  
 XX AC AAP20198;  
 XX DT 14-AUG-1992 (first entry)  
 XX DE Sequence of somatostatin-25 analogue.  
 XX KW Somatostatin; hormone; growth hormone release; inhibition.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /label= H-S  
 FT Disulfide-bond 14..25  
 FT Modified-site 25 /label= C-OH  
 XX US4316891-A.  
 XX 23-FEB-1982.  
 XX 14-JUN-1980; 80US-0159801.  
 XX 14-JUN-1980; 80US-0159801.  
 XX (SALK-) SALK INST BIOLOG.  
 XX Guillemin RCL, Esch FS, Bohlen P, Brazeau PE, Ling NC;  
 XX WPI; 1982-19801E/10.  
 XX Extended somatostatin analogues - with increased inhibition of  
 FT growth hormone release  
 XX Claim 6; Column 12; 7pp; English.  
 XX The inventors claim a pharmaceutical compsn. which comprises (i)  
 CC synthetic somatostatin-28 (SS-28), SS-25 or (D-Trip(22))-SS-28 and  
 CC (ii) a liq. or solid carrier, and SS-28 derivs. and SS-25 derivs.  
 CC The compsn. and derivs. are more potent than somatostatin in  
 CC inhibiting release of growth hormone; they also inhibit basal and  
 CC stimulating insulin and glucagon secretion. (D-Trip(22))-SS-28  
 CC exhibits very substantial increases in potency w.r.t. inhibition of  
 CC growth hormone secretion.

XX SQ Sequence 25 AA;  
 Query Match 22.4%; Score 127; DB 3; Length 25;  
 Best Local Similarity 91.3%; Pred. No. 3e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 92 PVLAPRRKAGCKNFFWKTFTSC 114  
 Db 3 pamaprerkagcknffwktftsc 25  
 RESULT 8  
 AAB91017  
 ID AAB91017 standard; Peptide; 25 AA.  
 XX AC AAB91017;  
 XX DT 22-JUN-2001 (first entry)  
 XX DE Somatostatin related peptide SEQ ID NO:191.  
 XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidyl; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200069900-A2.  
 XX 23-NOV-2000.  
 XX 17-MAY-2000; 2000WO-US13576.  
 XX 17-MAY-1999; 98US-0134406.  
 XX 10-SEP-1999; 98US-0153406.  
 XX 15-OCT-1999; 99US-0159783.  
 XX (CONJ-) CONJUCHEM INC.  
 XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 XX WPI; 2001-112059/12.  
 XX Modifying and attaching therapeutic peptides to albumin prevents  
 XX peptidase degradation, useful for increasing length of in vivo activity  
 XX Disclosure; Page 252; 733pp; English.  
 XX The present invention describes a modified therapeutic peptide (I)  
 XX comprising a therapeutically active amino acid region (III) and a  
 XX reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
 XX a less therapeutically active amino acid region (IV), which covalently  
 XX bonds with amino/hydroxyl/thiol groups on blood components to form a  
 XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 XX factors and neurotransmitters, to protect them from peptidase activity  
 XX in vivo for the treatment of various disorders. Endogenous therapeutic  
 XX peptides are not suitable as drug candidates as they require frequent  
 XX administration due to rapid degradation by peptidases in the body.  
 XX Modifying and attaching therapeutic peptides to albumin prevents or  
 XX reduces the action of peptidases to increase length of activity (half  
 XX life) and specificity as bonding to large molecules decreases  
 XX intracellular uptake and interference with physiological processes.  
 XX AAB90829 to AAB92441 represent peptides which can be used in the  
 XX exemplification of the present invention.  
 XX SQ Sequence 25 AA;

Query Match 22.4%; Score 127; DB 22; Length 25;  
 Best Local Similarity 91.3%; Pred. No. 3e-07;  
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114  
 Db 3 pamaprerkagcknffwktftsc 25

RESULT 9  
 AAW51859  
 ID AAW51859 standard; peptide; 28 AA.  
 AC AAW51859;  
 DT 10-SEP-1998 (first entry)  
 DE Mouse somatostatin SS-28 hormone.  
 KW Mouse somatostatin SS-28 hormone; growth hormone; insulin; glucagon;  
 KW thyroid stimulating hormone; octreotide; cell-based delivery of insulin;  
 KW glucose-stimulated insulin secretion; SSTRV; somatostatin receptor;  
 KW mouse somatostatin receptor type V gene; diabetes.  
 OS Mus musculus.  
 PN WO9935242-A1.  
 PD 15-JUL-1999.  
 PF 11-JAN-1999; 99WO-US00633.  
 PR 03-JUN-1998; 98US-0087848.  
 PR 12-JAN-1998; 98US-0071193.  
 PR 12-JAN-1998; 98US-0071209.  
 PR 12-JAN-1998; 98US-0072556.  
 PR 03-JUN-1998; 98US-0087821.  
 XX (BETA-) BETAGENE INC.  
 XX Clark SA, Quade C;  
 PI WPI; 1999-444195/37.  
 DR New defined medium for culture of neuroendocrine cells, e.g. of  
 PT insulin-secreting cells  
 XX Example 8; Page 143; 312pp; English.  
 XX The present sequence is a mouse somatostatin (SS-28) hormone which  
 CC was found to inhibit the release of growth hormone, thyroid stimulating  
 CC hormone, insulin and glucagon. In addition, SS-28 and its analogue  
 CC Octreotide may inhibit growth of some tumours. The hormone was  
 CC used to study its effect on glucose-stimulated insulin secretion in  
 CC high expressing and non-expressing clones of mouse somatostatin  
 CC receptor, type V gene (SSTRV). The insulin secretion was highly inhibited  
 CC in the high expressing clone as compared to the non-expressing clone  
 CC because the high expressing clone showed high sensitivity to  
 CC somatostatin. The hormone effectively inhibits insulin secretion in the  
 CC absence of glucose. The somatostatin receptor can be introduced in cell  
 CC lines, used in cell-based delivery of insulin for treating diabetes, for  
 CC precise regulation of insulin release.  
 XX Sequence 28 AA;  
 SQ

Query Match 22.4%; Score 127; DB 19; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-07;  
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114  
 Db 6 pamaprerkagcknffwktftsc 28

RESULT 11  
 AAW24384  
 ID AAW24384 standard; peptide; 28 AA.  
 XX AAW24384;  
 AC AAW24384;  
 DT 20-SEP-1999 (first entry)  
 XX

AAV28703  
 ID AAV28703 standard; peptide; 28 AA.  
 XX AAV28703;  
 DT 07-OCT-1999 (first entry)  
 DE Mouse somatostatin SS-28 hormone.  
 KW Mouse somatostatin SS-28 hormone; growth hormone; insulin; glucagon;  
 KW thyroid stimulating hormone; octreotide; cell-based delivery of insulin;  
 KW glucose-stimulated insulin secretion; SSTRV; somatostatin receptor;  
 KW mouse somatostatin receptor type V gene; diabetes.  
 OS Mus musculus.  
 PN WO9935242-A1.  
 PD 15-JUL-1999.  
 PF 11-JAN-1999; 99WO-US00633.  
 PR 03-JUN-1998; 98US-0087848.  
 PR 12-JAN-1998; 98US-0071193.  
 PR 12-JAN-1998; 98US-0071209.  
 PR 12-JAN-1998; 98US-0072556.  
 PR 03-JUN-1998; 98US-0087821.  
 XX (BETA-) BETAGENE INC.  
 XX Clark SA, Quade C;  
 PI WPI; 1999-444195/37.  
 DR New defined medium for culture of neuroendocrine cells, e.g. of  
 PT insulin-secreting cells  
 XX Example 8; Page 143; 312pp; English.  
 XX The present sequence is a mouse somatostatin (SS-28) hormone which  
 CC was found to inhibit the release of growth hormone, thyroid stimulating  
 CC hormone, insulin and glucagon. In addition, SS-28 and its analogue  
 CC Octreotide may inhibit growth of some tumours. The hormone was  
 CC used to study its effect on glucose-stimulated insulin secretion in  
 CC high expressing and non-expressing clones of mouse somatostatin  
 CC receptor, type V gene (SSTRV). The insulin secretion was highly inhibited  
 CC in the high expressing clone as compared to the non-expressing clone  
 CC because the high expressing clone showed high sensitivity to  
 CC somatostatin. The hormone effectively inhibits insulin secretion in the  
 CC absence of glucose. The somatostatin receptor can be introduced in cell  
 CC lines, used in cell-based delivery of insulin for treating diabetes, for  
 CC precise regulation of insulin release.  
 XX Sequence 28 AA;  
 SQ

Query Match 22.4%; Score 127; DB 20; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114  
 Db 6 pamaprerkagcknffwktftsc 28

RESULT 11  
 AAW24384  
 ID AAW24384 standard; peptide; 28 AA.  
 XX AAW24384;  
 AC AAW24384;  
 DT 20-SEP-1999 (first entry)  
 XX

Query Match 22.4%; Score 127; DB 19; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114  
 Db 6 pamaprerkagcknffwktftsc 28

RESULT 10

Query Match 22.4%; Score 127; DB 20; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 92 PVLAPRRKAGCKNFFWKTFTSC 114  
 Db 6 pamaprerkagcknffwktftsc 28

RESULT 11  
 AAW24384  
 ID AAW24384 standard; peptide; 28 AA.  
 XX AAW24384;  
 AC AAW24384;  
 DT 20-SEP-1999 (first entry)  
 XX



XX PD 23-NOV-2000.  
 XX PF 17-MAY-2000; 2000WO-US13576.  
 XX PR 17-MAY-1999; 99US-0134406.  
 XX PR 10-SEP-1999; 99US-0153406.  
 XX PR 15-OCT-1999; 99US-0159783.  
 XX PA (CONJ-) CONJUCHEM INC.  
 XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 XX DR WPI; 2001-112059/12.  
 XX PT Modifying and attaching therapeutic peptides to albumin prevents  
 XX PT peptidase degradation, useful for increasing length of in vivo activity  
 XX PS  
 XX  
 XX Disclosure; Page 252; 733pp; English.

CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimide and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the activity of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.

XX Sequence 28 AA;

Query Match 22.4%; Score 127; DB 22; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 92 PVLAPRRKAGCKNFFWKFTTSC 114  
 | :|||||  
 6 pamaprerkagcknffwkftsc 28

RESULT 14  
 AAU07669  
 ID AAU07669 standard; Peptide; 28 AA.  
 XX AC AAU07669;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE Mammalian somatostatin 28. (SS-28) peptide sequence.

XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;  
 KW PPSS-II; PPSS-III; endocrine tumour; pituitary gland; glucagonoma; AIDS;  
 KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;  
 KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone; SS-28;  
 KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;  
 KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;  
 KW neoplasm; metastasis; gene therapy; antidiabetic; neurotropic; cytostatic;  
 KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.

XX Mammalia sp.  
 XX CA2325169-A1.

XX PD 03-JUN-2001.  
 XX PF 01-DEC-2000; 2000CA-2325169.  
 XX PR 03-DEC-1999; 99US-0168934.  
 XX PA (NDSU-) NDSU RES FOUND.  
 XX PI Sheridan MA, Moore CA, Kittelson JD;  
 XX DR WPI; 2001-425997/46.  
 XX PT New somatostatin polypeptides derived from Oncorhynchus mykiss, useful  
 XX PT for treating diabetes mellitus, acromegaly, gastrinoma, acquired  
 XX PT immunodeficiency syndrome and neurological disorders -  
 XX PS Example 5; Fig 6; 52pp; English.

CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide  
 CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of  
 CC preprosomatostatin II (PPSS-II). The protein sequences and their  
 CC associated polynucleotides are useful for identifying modified  
 CC somatostatin polypeptides which functions as a somatostatin agonist useful  
 CC for research, therapeutics or diagnostics, including medical and  
 CC veterinary applications. The wild-type somatostatin and its modified  
 CC version are useful for treating hypersecretion from endocrine tumours in  
 CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.  
 CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage  
 CC through their effects on cell proliferation and apoptosis and as adjuncts  
 CC in the treatment of diabetes mellitus via inhibition of growth hormone  
 CC and glucagon. In addition, dysfunctional somatostatin secretion is  
 CC associated with acquired immunodeficiency syndrome (AIDS) and various  
 CC neurological disorders (e.g. epilepsy, Alzheimer's disease and  
 CC Huntington's disease) and somatostatin antagonists are effective in the  
 CC treatment of such conditions. Nucleic acids encoding the polypeptides are  
 CC useful in gene therapy and fusion peptides can be targeted to neoplasms  
 CC and their metastases, inhibiting the release of their secretory products.  
 CC This sequence represents the mammalian somatostatin 28 (SS-28) peptide.

XX Sequence 28 AA;

Query Match 22.4%; Score 127; DB 22; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 92 PVLAPRRKAGCKNFFWKFTTSC 114  
 | :|||||  
 6 pamaprerkagcknffwkftsc 28

RESULT 15  
 AAB91020  
 ID AAB91020 standard; Peptide; 29 AA.  
 XX AC AAB91020;  
 XX DT 22-JUN-2001 (first entry)  
 XX DE Somatostatin related peptide SEQ ID NO:194.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimide; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.  
 XX Synthetic.  
 XX WO200069900-A2.  
 XX PN 23-NOV-2000.

Fri Jun 14 08:03:53 2002

PF 17-MAY-2000; 2000WO-US13576.  
 XX  
 PR 17-MAY-1999; 99US-0134406.  
 PR 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0159783.  
 XX  
 PA (CONJ-) CONJUCHEM INC.  
 XX  
 PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 XX  
 DR WPI; 2001-112059/12.  
 XX  
 PT Modifying and attaching therapeutic peptides to albumin prevents  
 PT peptidase degradation, useful for increasing length of in vivo activity  
 PT  
 XX  
 PS Disclosure; Page 253; 733pp; English.  
 XX  
 CC The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 29 AA;

Query Match 22.4%; Score 127; DB 22; Length 29;  
 Best Local Similarity 91.3%; Pred. NO. 3.7e-07;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 92 PVLAPRERKAGCKNFFWKFTSC 114  
 Db 7 pamaprerkagcknffwkftsc 29

Search completed: June 13, 2002, 12:22:20  
 Job time: 273 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw.model

Run on: June 13, 2002, 12:19:42 ; Search time 28.68 Seconds  
(without alignments)  
97.089 Million cell updates/sec

Title: US-09-727-739B-3  
Perfect score: 566  
Sequence: 1 MLSTRVOCALALLSLALALS.....APRRKACGRNFFWKTFTSC 114

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cnp2.6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cnp2.6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cnp2.6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cnp2.6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cnp2.6/ptodata/2/1aa/PTCUS.COMB.pep.\*  
6: /cnp2.6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	329	58.1	110	3	US-08-648-322-3
2	94.5	16.7	112	3	US-08-648-322-2
3	94.5	16.7	112	4	US-09-001-472-3
4	92	16.3	943	1	US-08-455-970A-10
5	87	15.4	14	1	US-07-977-628A-1
6	87	15.4	14	1	US-08-255-272-5
7	87	15.4	14	1	US-08-416-007-4
8	87	15.4	14	1	US-08-676-263-11
9	87	15.4	14	1	US-08-286-748B-13
10	87	15.4	14	1	US-08-690-090A-1
11	87	15.4	14	2	US-08-488-159-1
12	87	15.4	14	2	US-08-453-764-1
13	87	15.4	14	2	US-08-475-751-4
14	87	15.4	14	2	US-08-282-980B-1
15	87	15.4	14	2	US-08-747-137-13
16	87	15.4	14	2	US-09-039-062-1
17	87	15.4	14	2	US-09-042-224-1
18	87	15.4	14	2	US-09-042-315A-1
19	87	15.4	14	3	US-08-931-095-1
20	87	15.4	14	3	US-09-100-414B-83
21	87	15.4	14	4	US-09-420-866-1
22	87	15.4	14	4	US-09-420-865-1
23	87	15.4	14	4	US-09-303-323-83
24	87	15.4	14	4	US-08-586-670A-1
25	87	15.4	14	4	US-09-120-237-1
26	87	15.4	14	6	5212156-1
27	87	15.4	14	6	5268278-4

28	87	15.4	15	6	5268278-1	Patent No. 5268278
29	87	15.4	31	3	US-09-100-414B-84	Sequence 84, Appl
30	87	15.4	31	3	US-09-100-414B-85	Sequence 85, Appl
31	87	15.4	31	3	US-09-100-414B-86	Sequence 86, Appl
32	87	15.4	31	3	US-09-100-414B-87	Sequence 87, Appl
33	87	15.4	31	3	US-09-100-414B-88	Sequence 88, Appl
34	87	15.4	31	3	US-09-100-414B-89	Sequence 89, Appl
35	87	15.4	31	4	US-09-303-323-84	Sequence 84, Appl
36	87	15.4	31	4	US-09-303-323-85	Sequence 85, Appl
37	87	15.4	31	4	US-09-303-323-86	Sequence 86, Appl
38	87	15.4	31	4	US-09-303-323-87	Sequence 87, Appl
39	87	15.4	31	4	US-09-303-323-88	Sequence 88, Appl
40	87	15.4	31	4	US-09-303-323-89	Sequence 89, Appl
41	87	15.4	34	3	US-09-100-414B-90	Sequence 90, Appl
42	87	15.4	34	4	US-09-303-323-90	Sequence 90, Appl
43	86.5	15.3	105	4	US-09-001-472-2	Sequence 2, Appl
44	85	15.0	89	3	US-08-648-322-11	Sequence 11, Appl
45	85	15.0	84	3	US-08-648-322-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-08-648-322-3  
; Sequence 3, Application US/08648322  
; Patent No. 6074872  
; GENERAL INFORMATION:  
; APPLICANT: Sutcliffe, Gregor J.  
; APPLICANT: de Lecea, Luis  
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE SCRIPTS RESEARCH INSTITUTE  
; STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/648,322  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: 519.0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 554-6312  
; TELEFAX: (619) 554-6312  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 110 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: C-terminal  
; US-08-648-322-3

Query Match 58.1%; Score 329; DB 3; Length 110;  
Best Local Similarity 59.1%; Pred. No. 5e-32;  
Matches 65; Conservative 16; Mismatches 27; Indels 1;  
Gaps 1;  
QY 7 QCALALLSLALSSVSAAPSDAKRLQRLQSLRMAPAGKQELARNLTVLLSELAHVENE 66  
DB 1 QCALANICVILAGVGTGASDPRLQFLQKSLAATGKQELAKYFLAELLSEPNQTFEND 60



;; FILING DATE: 31-MAY-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/960,932  
;; FILING DATE: 14-OCT-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: ROBINS, ROBERTA L.  
;; REGISTRATION NUMBER: 33,208  
;; REFERENCE/DOCKET NUMBER: 9001-0016.10  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 327-3400  
;; TELEFAX: (415) 327-3231  
;; INFORMATION FOR SEQ ID NO: 10:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 943 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-455-970A-10

Query Match 16.3%; Score 92; DB 1; Length 943;  
Best Local Similarity 30.7%; Pred. No. 0.021;  
Matches 31; Conservative 5; Mismatches 23; Indels 42; Gaps 3;

QY 14 SLALATSSVAAPSADAKLQRLQSLMAPAGKQELARNTLVLLSELARHVENAEIELDDM 73

Db 885 SLDKLSSVSFTSSNDSRNV---LVAPT-----SMLDQSLSSLOFARGSS----- 927

QY 74 SHGVEQEDVLELRAPGVLPAPRRKAGCKNFFWKTFTSC 114

Db 928 -----SSAGCKNFFWKTFTSC 943

## RESULT 5

US-07-977-628A-1.  
Sequence 1, Application US/07977628A  
Patent No. 5405597

## GENERAL INFORMATION:

APPLICANT: Dean, Richard T  
APPLICANT: Lister-James, John  
APPLICANT: Buttram, Scott  
TITLE OF INVENTION: Technetium-99m Labeled Somatostatin-  
DERIVED PEPTIDES FOR IMAGING  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07977,628A  
FILING DATE: 17-NOV-1992

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5405597nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 91,642-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid

;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 2..14  
;; OTHER INFORMATION: /label- Disulfide-bond  
;; OTHER INFORMATION: /note- "The sidechain thiol groups of the third  
;; OTHER INFORMATION: residue-cysteine and the carboxy-terminal cysteine  
;; OTHER INFORMATION: form a disulfide bond in native somatostatin  
;; US-07-977-628A-1

Query Match 15.4%; Score 87; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKTFTSC 114

Db 1 AGCKNFFWKTFTSC 14

## RESULT 6

US-08-255-272-5  
Sequence 5, Application US/08255272  
Patent No. 5627268

## GENERAL INFORMATION:

APPLICANT: Kumar, Ramesh  
APPLICANT: Sharma, Ajay  
APPLICANT: Khoury-Christianson, Anastasia  
APPLICANT: M  
TITLE OF INVENTION: Production of Therapeutic Peptides in  
TITLE OF INVENTION: Transgenic Animals as a Fusion with Hemoglobin  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2711

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/255,272

## FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30742  
REFERENCE/DOCKET NUMBER: 6794-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid

## TOPOLOGY:

MOLECULE TYPE: unknown  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-255-272-5

Query Match 15.4%; Score 87; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKTFTSC 114

Db 1 AGCKNFFWKTFTSC 14



DB 1 AGCKNFFWKFTSC 14

## RESULT 7

US-08-416-007-4  
; Sequence 4, Application US/08416007  
; Patent No. 5693679  
; GENERAL INFORMATION:  
; APPLICANT: Vincent, Jean-Pierre  
; APPLICANT: Gaudinault, Georges  
; APPLICANT: Beaudet, Alain  
; TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/416,007  
; FILING DATE: 04-APR-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 06942/003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-416-007-4

Query Match 15.4%; Score 87; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTSC 114

DB 1 AGCKNFFWKFTSC 14

## RESULT 8

US-08-676-263-11  
; Sequence 11, Application US/08676263  
; Patent No. 5705143  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Gary R.  
; APPLICANT: Forster, Alan M.  
; APPLICANT: Riley, Anthony L. M.  
; APPLICANT: Storey, Anthony E.  
; TITLE OF INVENTION: BIOLOGICAL TARGETING AGENTS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America

ZIP: 60606-6402  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/676,263  
; FILING DATE: 07-NOV-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 94300224.6  
; FILING DATE: 12-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: (312) 474-6600  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Cross-links  
; LOCATION: 3..14  
US-08-676-263-11

Query Match 15.4%; Score 87; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTSC 114

DB 1 AGCKNFFWKFTSC 14

## RESULT 9

US-08-286-748B-13  
; Sequence 13, Application US/08286748B  
; Patent No. 5759342  
; GENERAL INFORMATION:  
; APPLICANT: Victor Gurewich  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3 5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 58SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: Wordperfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,748B  
; FILING DATE: August 5, 1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J. Peter Fasse

REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04547/013001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-286-748B-13

Query Match 15.4%; Score 87; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 101 AGCKNFFWKTFTSC 114  
|||||

-Db 1 AGCKNFFWKTFTSC 14

RESULT 10  
US-08-690-090A-1  
Sequence 1, Application US/08690090A  
Patent No. 5770687  
GENERAL INFORMATION:  
APPLICANT: HORNIK, VERED  
APPLICANT: SERT-LEVY, ALON  
APPLICANT: GELLERMAN, GARY  
APPLICANT: GILON, CHAIM  
TITLE OF INVENTION: Conformationally Constrained Backbone  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/690,090A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/488,159  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fanucci, Allan A.  
REGISTRATION NUMBER: 30,256  
REFERENCE/DOCKET NUMBER: 7754-052-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-690-090A-1

Query Match 15.4%; Score 87; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKTFTSC 114  
|||||

Db 1 AGCKNFFWKTFTSC 14

RESULT 11  
US-08-488-159-1  
Sequence 1, Application US/08488159  
Patent No. 5811392  
GENERAL INFORMATION:  
APPLICANT: Gilson, Chaim  
TITLE OF INVENTION: Conformationally Constrained Backbone  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,159  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fanucci, Allan A.  
REGISTRATION NUMBER: 30,256  
REFERENCE/DOCKET NUMBER: 7754-033  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-488-159-1

Query Match 15.4%; Score 87; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKTFTSC 114  
|||||

Db 1 AGCKNFFWKTFTSC 14

RESULT 12  
US-08-465-764-1  
Sequence 1, Application US/08465764  
Patent No. 5814298  
GENERAL INFORMATION:  
APPLICANT: Dean, Richard T.  
APPLICANT: Lister-James, John  
TITLE OF INVENTION: Technetium-99m Labeled  
Somatostatin-derived Peptides for Imaging and Therapeutic  
Uses  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,764  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5814298nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 92,385-R  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-3317  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 3..14  
OTHER INFORMATION: /label= Disulfide bond  
OTHER INFORMATION: /note= "The peptide is cyclized between the  
sidechain sulfur atoms of the 3d and 14th residues  
US-08-465-764-1

Query Match 15.4%; Score 87; DB 2; Length 14;  
Best Local Similarity 100.0%; Fred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ACCKNFFWKTFTSC 114  
Db 1 ACCKNFFWKTFTSC 14  
|||||

---RESULT 13  
1-08-475-751-4  
Sequence 4, Application US/08475751  
Patent No. 5824772  
GENERAL INFORMATION:  
APPLICANT: Vincent, Jean-Pierre  
APPLICANT: Gaudriault, Georges  
APPLICANT: Beaudet, Alain  
TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 585 Commercial Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-1024  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,751  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/416,007  
FILING DATE: 04-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06942/004001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/723-4123  
TELEFAX: 617/723-8962  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-475-751-4

Query Match 15.4%; Score 87; DB 2; Length 14;  
Best Local Similarity 100.0%; Fred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ACCKNFFWKTFTSC 114  
Db 1 ACCKNFFWKTFTSC 14  
|||||

---RESULT 14  
US-08-282-980B-1  
Sequence 1, Application US/08282980B  
Patent No. 5932189  
GENERAL INFORMATION:  
APPLICANT: Dean, Richard T.  
APPLICANT: McBride, William  
APPLICANT: Lister-James, John  
TITLE OF INVENTION: Peptides  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff.  
STREET: 300 South Wacker Drive, Seventh Floor  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,980B  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5932189nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 92,385-I  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 3..14

OTHER INFORMATION: /label- Disulfide bond  
OTHER INFORMATION: /note- "A disulfide bond exists between the  
OTHER INFORMATION: two sulfur atoms of the cysteine residues.  
US-08-282-980B-1

Query Match 15.4%; Score 87; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114  
|||||  
DB 1 AGCKNFFWKFTTSC 14

## RESULT 15

US-08-747-137-13  
; Sequence 13, Application US/08747137  
; Patent No. 5945033

## GENERAL INFORMATION:

APPLICANT: YEN, Richard C.K.  
TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR  
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE

NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA

COUNTRY: USA  
ZIP: 94111

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/747,137

FILING DATE: 12-NOV-1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,546

FILING DATE: 14-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US-08/069,831

FILING DATE: 01-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/959,560

FILING DATE: 13-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/641,720

FILING DATE: 15-JAN-1991

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 016197-000840US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-576-0200

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

US-08-747-137-13

Query Match 15.4%; Score 87; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 AGCKNFFWKFTTSC 114  
|||||

DB 1 AGCKNFFWKFTTSC 14

Search completed: June 13, 2002, 12:22:55  
Job time: 193 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:20:22 ; Search time 35.79 Seconds  
(without alignments)  
306.068 Million cell updates/sec

Title: US-09-727-739b-3  
Perfect score: 566  
Sequence: 1 MLSTRVQCALALLSLALAISSVRAAPS...APRRKAGCKNFFWTFISC 114

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

otal number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR-71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	409	72.3	114	1 RIIDS1	somatostatin-14 pr
2	409	72.3	114	2 IS0798	preprosomatostatin
3	400	70.7	116	1 S20630	somatostatin precu
4	382.5	67.6	115	2 JC6166	somatostatin-14 pr
5	360	63.6	116	1 RIHUS1	somatostatin I pre
6	360	63.6	116	1 A28968	somatostatin I pre
7	352	62.2	116	1 RIHOS1	somatostatin precu
8	347	61.3	116	1 RIHPS1	somatostatin precu
9	347	61.3	116	1 RIHPS1	somatostatin precu
10	293	51.8	92	1 RIPOS1	somatostatin I pre
11	251	44.3	121	1 RIAPS1	somatostatin I pre
12	177.5	31.4	115	2 IS1064	somatostatin II pr
13	162.5	28.7	125	1 RIAPS2	somatostatin II pr
14	127	22.4	28	2 A61322	somatostatin-28 -
15	116.5	20.6	34	2 A32271	somatostatin-relat
16	114	20.1	73	2 S00169	somatostatin II pr
17	110.5	19.5	105	1 RIIDS2	somatostatin-22 pr
18	107.5	19.0	103	2 JC6167	somatostatin-14 [p
19	107	18.9	25	2 A60842	somatostatin-25 -
20	106	18.7	37	2 A32000	somatostatin, panc
21	103.5	18.3	74	2 S00166	somatostatin II pr
22	94.5	16.7	112	2 S67489	cortistatin precu
23	87	15.4	14	2 C60414	somatostatin - sll
24	87	15.4	14	2 B60842	somatostatin I - C
25	87	15.4	14	2 A60840	somatostatin I - C
26	87	15.4	14	2 S00172	somatostatin I - s
27	87	15.4	25	2 B60840	somatostatin-25 -
28	86.5	15.3	105	2 JC5414	cortistatin-like p
29	82	14.5	14	2 A60622	somatostatin - spo

30	79.5	14.0	846	2 SS2418	GTP-binding regula
31	77.5	13.7	571	2 AE3281	hypothetical cytos
32	77	13.6	385	2 T18621	hypothetical prote
33	74	13.1	297	2 T48855	probable heat shoc
34	74	13.1	965	2 AE0418	valine--trna ligas
35	74	13.1	1385	2 T13887	tir protein - frul
36	74	13.1	1389	2 T13852	gene wheeler prote
37	74	13.1	1732	2 T14039	protein kinase (EC
38	73	12.9	493	2 S49175	legumain (EC 3.4.2
39	72.5	12.8	216	2 T40268	ubiquitinone biosynt
40	72.5	12.8	274	2 S72643	type 4 fibrial bi
41	71.5	12.6	1285	2 H85041	hypothetical prote
42	71	12.5	1295	2 T24587	hypothetical prote
43	70.5	12.5	478	2 AB3641	fumarate hydratase
44	70.5	12.5	925	2 E83529	sensor/response re
45	70	12.4	414	2 T15947	hypothetical prote

## ALIGNMENTS

### RESULT 1

RIIDS1  
somatostatin-14 precursor - channel catfish  
N;Alternate names: somatostatin I  
N;Contains: somatostatin-14  
C;Species: Ictalurus punctatus (Channel catfish)  
C;Date: 30-Jun-1980 #sequence,revision 31-Dec-1993 #text\_change 18-Jun-1999  
C;Accession: S00292; A93897; A92334; A01435  
R;Minth, C.D.; Taylor, W.L.; Magazin, M.; Tavlanini, M.A.; Collier, K.; Weith, H.L.;  
J. Biol. Chem. 257, 10372-10377, 1982  
A;Title: The structure of cloned DNA complementary to catfish pancreatic somatostati  
A;Reference number: S00292; MUID:82265698  
A;Accession: S00292  
A;Molecule type: mRNA  
A;Residues: 1-114 <MIN>  
A;Cross-references: EMBL:V00607; NID:964017; PIDN:CAA23877.1; PID:964018  
R;Taylor, W.L.; Collier, K.J.; Deschenes, R.J.; Weith, H.L.; Dixon, J.E.  
Proc. Natl. Acad. Sci. U.S.A. 78, 6694-6698, 1981  
A;Title: Sequence analysis of a cDNA coding for a pancreatic precursor to somatostat  
A;Reference number: A93897; MUID:82082515  
A;Accession: A93897  
A;Molecule type: mRNA  
A;Residues: 82-108 <TAY>  
A;Cross-references: GB:J00944  
R;Andrews, P.C.; Dixon, J.E.  
J. Biol. Chem. 256, 8267-8270, 1981  
A;Title: Isolation and structure of a peptide hormone predicted from a mRNA sequence  
A;Reference number: A92334; MUID:81264223  
A;Accession: A92334  
A;Molecule type: protein  
A;Residues: 101-114 <END>  
C;Superfamily: somatostatin  
C;Keywords: neuropeptide  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-100/Domain: propeptide #status predicted <PRO>  
F;101-114/Product: somatostatin-14 #status experimental <MAT>  
F;103-114/Disulfide bonds: #status experimental

Query Match 72.3% Score 409; DB 1; Length 114;  
Best Local Smilarity 69.3% Pred. No. 4.5e-32;  
Matches 79; Conservative 15; Mismatches 20; Indels 0; Gaps 0;

Qy	1	MLSTRVQCALALLSLALAISSVRAAPS	SDAKLROLLOSLMAPAGKQELARNTLVLSSEL	60
Db	1	MPSTRVQCALALLSLALAISSVRAAPS	SDAKLROLLOSLMAPAGKQELARNTLVLSSEL	60
Qy	61	AHVENEAIELDDNMGVQEQVDVLELER	APCPVLAPRRKAGCKNFFWTFISC	114
Db	61	AQAEENVLDSDVSRRAESGARLEMER	APCPVLAPRRKAGCKNFFWTFISC	114

## RESULT 2

preprosomatostatin SS-14 - channel catfish  
 C:Species: Ictalurus punctatus (Channel catfish)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
 C:Accession: J50798  
 R:Dixon, J.E.; Andrews, P.C.  
 Adv. Exp. Med. Biol. 188, 19-29, 1985  
 A:Title: Somatostatins of the channel catfish.  
 A:Reference number: J50798; MUID:85303576  
 A:Accession: J50798  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-114 <DIX>  
 A:Cross-references: GB:M25903; NID:g213339; PIDN:AAA49339.1; PID:g213340  
 C:Superfamily: somatostatin

Query Match 72.3%; Score 409; DB 2; Length 114;  
 Best Local Similarity 69.3%; Pred. No. 4.5e-32;  
 Matches 79; Conservative 15; Mismatches 20; Indels 0; Gaps 0;  
 QY 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQRLSRLMAPAGKQELARNTLVLLSEL 60  
 DB 1 MSTRICQALLAVALSVCSVSGAPSDAKLRQLLQRLSRLMAPAGKQELARNTLVLLSEL 60  
 QY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PGPVLPAPRRKAGCKNFFWKTFTSC 114  
 DB 61 ABAENEVLDSVRAASEGARLEMERAGPMLAPRRKAGCKNFFWKTFTSC 114

## RESULT 3

S20630  
 somatostatin precursor - chicken  
 C:Species: Gallus gallus (Chicken)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 R:Nata, K.; Kobayashi, T.; Karabashi, K.; Kato, S.; Yamamoto, H.; Okamoto, M.  
 submitted to the EMBL Data Library, June 1991  
 A:Description: Nucleotide sequence determination of chicken somatostatin precursor cDNA.  
 A:Reference number: S20630  
 A:Accession: S20630  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <NAT>  
 A:Cross-references: EMBL:X60191; NID:g62985; PIDN:CAA42747.1; PID:g62986  
 C:Superfamily: somatostatin

Query Match 70.7%; Score 400; DB 1; Length 116;  
 Best Local Similarity 69.8%; Pred. No. 3.4e-31;  
 Matches 81; Conservative 14; Mismatches 19; Indels 2; Gaps 1;  
 QY 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQRLSRLMAPAGKQELARNTLVLLSEL 60  
 DB 1 MLSCFQCALALLSLALAVGTVAAPSDPRLRQFLQKSLAAAGKQELAKYFLAELLSEP 60  
 QY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PGPVLPAPRRKAGCKNFFWKTFTSC 114  
 DB 61 SOTENEAESEDLRSRGAQEDEVRLERSANSNPALAPRRKAGCKNFFWKTFTSC 116

## RESULT 4

JC6166  
 somatostatin-14 precursor - laughing frog  
 N:Alternate names: PSS1 protein  
 C:Species: Rana ridibunda (laughing frog)  
 C:Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 16-Jul-1999  
 C:Accession: JC6166  
 R:Tostivint, H.; Lihmann, I.; Buchares, C.; Vileau, D.; Coulouarn, Y.; Fournier, A.; C  
 Proc. Natl. Acad. Sci. U.S.A. 93, 12605-12610, 1996  
 A:Title: Occurrence of two somatostatin variants in the frog brain: Characterization of  
 A:Reference number: JC6166; MUID:97057290

## A:Contents: brain

A:Accession: JC6166  
 A:Molecule type: mRNA  
 A:Residues: 1-115 <TOS>  
 A:Cross-references: GB:U68136; NID:g1890650; PIDN:AAC60093.1; PID:g1890651  
 C:Comment: This protein acts both as a neurotransmitter/neuromodulator and a hormone.  
 C:Genetics:  
 A:Gene: pss1  
 C:Superfamily: somatostatin  
 C:Keywords: brain; hormone

Query Match 67.6%; Score 382.5; DB 2; Length 115;  
 Best Local Similarity 69.0%; Pred. No. 1.6e-29;  
 Matches 80; Conservative 13; Mismatches 20; Indels 3; Gaps 2;  
 QY 1 MLSTRVOCALALLSLALAISSVSAAPSDAKLRQLLQRLSRLMAPAGKQELARNTLVLLSEL 60  
 DB 1 MOSCRVOCALTLTLLSALAINISAAPDPLRQFLQKSL-ASAGKQELAKYFLAELLSEP 59  
 QY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PGPVLPAPRRKAGCKNFFWKTFTSC 114  
 DB 60 SOTENEAESEDLPRGAQEDEVRLERSANSNPALAPRRKAGCKNFFWKTFTSC 115

## RESULT 5

RHUS1  
 somatostatin I precursor - human  
 N:Alternate names: preprosomatostatin  
 C:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 18-Jun-1999  
 C:Accession: A43614; A01430; S09381; S50024  
 R:Shen, L.P.; Rutter, W.J.  
 Science 224, 168-171, 1984  
 A:Title: Sequence of the human somatostatin I gene.  
 A:Reference number: A43614; MUID:84146798  
 A:Accession: A43614  
 A:Molecule type: DNA  
 A:Residues: 1-116 <SH2>  
 A:Cross-references: GB:J00306; NID:g338287; PIDN:AAA60566.1; PID:g338288  
 R:Shen, L.P.; Pictet, R.L.; Rutter, W.J.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 4575-4579, 1982  
 A:Title: Human somatostatin I: sequence of the cDNA.  
 A:Reference number: A01430; MUID:83014931  
 A:Accession: A01430  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <SHE>  
 A:Cross-references: GB:J00306; NID:g338287; PIDN:AAA60566.1; PID:g338288  
 A:Experimental source: pancreatic somatostatinoma  
 R:Gomez, S.; Boileau, G.; Zollinger, L.; Nault, C.; Rholam, M.; Cohen, P.  
 EMBO J. 8, 2911-2916, 1989  
 A:Title: Site-specific mutagenesis identifies amino acid residues critical in prohorm

A:Accession: S09381; MUID:90059875  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 88-108 <GOM>  
 R:Odum, L.; Johnsen, A.H.  
 Biochem. J. 303, 263-268, 1994  
 A:Title: Human seminal plasma contains somatostatin-64.  
 A:Reference number: S50024; MUID:95031969  
 A:Accession: S50024  
 A:Status: preliminary  
 A:Molecule type: protein  
 C:Residues: 53-62; 67-82 <ODU>  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Genetics:  
 A:Gene: GDB:SST  
 A:Cross-references: GDB:119604; OMIM:182450  
 A:Map position: 3q28-3q28  
 A:Introns: 46/3  
 C:Function:

A:Description: inhibits the secretion of a number of peptide hormones, including somatostatin  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; hypothalamus; neuropeptide  
 F:25-88/Domain: signal sequence #status predicted <SIG>  
 F:59-116/Product: somatostatin-28 #status predicted <PRO>  
 F:103-116/Product: somatostatin-14 #status predicted <M28>  
 F:105-116/Disulfide bonds: #status experimental

Query Match 63.6%; Score 360; DB 1; Length 116;  
 Best Local Similarity 62.1%; Pred. No. 2.2e-27;  
 Matches 72; Conservative 17; Mismatches 25; Indels 2; Gaps 1;

OY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLROLLQRLSRLMAPAGKQELARNLTVELLSEL 60  
 DB 1 MLSCRLOCALAALSIVLALGCVTGAPSDPRFLQKSLAAAGKQELAKYFLAELLSEP 60  
 OY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PGPVLAPEPRKAGCKNFFWKTFTSC 114  
 b 61 NOTENDALEDPELSQAAEQDEMRLQRLSANSNPAMPAPRERKAGCKNFFWKTFTSC 116

RESULT  
 A28968

somatostatin I precursor - crab-eating macaque  
 N:Alternate names: preprosomatostatin  
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Macaca fascicularis (crab-eating macaque)  
 C:Date: 30-Jun-1989 #sequence\_revision 31-Jan-1997 #text\_change 18-Jun-1999  
 C:Accession: A28968  
 R:Travis, G.H.; Sutcliffe, J.G.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 1696-1700, 1988  
 A:Title: Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low-  
 A:Reference number: A28968; MUID:88144503  
 A:Accession: A28968  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <TRA>  
 A:Cross-references: GB:M19318; NID:g342298; PIDN:AAA36908.1; PID:g342299  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; neuropeptide  
 F:25-88/Domain: signal sequence #status predicted <SIG>  
 F:59-116/Product: somatostatin-28 #status predicted <PRO>  
 F:103-116/Product: somatostatin-14 #status predicted <M28>  
 F:105-116/Disulfide bonds: #status predicted

Query Match 63.6%; Score 360; DB 1; Length 116;  
 Best Local Similarity 62.1%; Pred. No. 2.2e-27;  
 Matches 72; Conservative 17; Mismatches 25; Indels 2; Gaps 1;

OY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLROLLQRLSRLMAPAGKQELARNLTVELLSEL 60  
 DB 1 MLSCRLOCALAALSIVLALGCVTGAPSDPRFLQKSLAAAGKQELAKYFLAELLSEP 60  
 OY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PGPVLAPEPRKAGCKNFFWKTFTSC 114  
 DB 61 NOTENDALEDPELSQAAEQDEMRLQRLSANSNPAMPAPRERKAGCKNFFWKTFTSC 116

RESULT  
 RIBOS1

somatostatin precursor - bovine  
 N:Alternate names: preprosomatostatin  
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 06-Mar-1992 #sequence\_revision 31-Jan-1997 #text\_change 18-Jun-1999  
 C:Accession: A40929  
 R:Su, C.J.; White, J.W.; Li, W.H.; Luo, C.C.; Frazier, M.L.; Saunders, G.F.; Chan, L.  
 Mol. Endocrinol. 2, 209-216, 1988  
 A:Title: Structure and evolution of somatostatin genes.

A:Reference number: A40929; MUID:88288237

A:Accession: A40929  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <SUA>  
 A:Cross-references: GB:M31217; NID:g163636; PIDN:AAA30744.1; PID:g163637  
 A:Note: the authors translated the codon ATT for residue 65 as Asn  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; neuropeptide  
 F:25-88/Domain: signal sequence #status predicted <SIG>  
 F:59-116/Product: somatostatin-28 #status predicted <PRO>  
 F:103-116/Product: somatostatin-14 #status predicted <M28>  
 F:105-116/Disulfide bonds: #status predicted

Query Match 62.2%; Score 352; DB 1; Length 116;  
 Best Local Similarity 61.2%; Pred. No. 1.3e-26;  
 Matches 71; Conservative 17; Mismatches 26; Indels 2; Gaps 1;

OY 1 MLSTRVQCALALLSLALAISSVSAAPSDAKLROLLQRLSRLMAPAGKQELARNLTVELLSEL 60  
 DB 1 MLSCRLOCALAALSIVLALGCVTGAPSDPRFLQKSLAAAGKQELAKYFLAELLSEP 60  
 OY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PGPVLAPEPRKAGCKNFFWKTFTSC 114  
 DB 61 NOTENDALEDPELSQAAEQDEMRLQRLSANSNPAMPAPRERKAGCKNFFWKTFTSC 116

RESULT 8

RIRTS1  
 somatostatin precursor - rat  
 N:Alternate names: preprosomatostatin  
 N:Contains: somatostatin-14; somatostatin-28  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 18-Jun-1999  
 C:Accession: A0983; A01431; A47598; A25229; I51829  
 R:Montminy, M.R.; Goodman, R.H.; Horovitch, S.J.; Habener, J.F.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 3337-3340, 1984  
 A:Title: Primary structure of the gene encoding rat preprosomatostatin.  
 A:Reference number: A20983; MUID:84221954  
 A:Accession: A20983  
 A:Molecule type: DNA  
 A:Residues: 1-116 <MON>  
 A:Cross-references: GB:J00787; NID:g207024; PIDN:AAA42164.1; PID:g207025  
 A:Note: the authors translated the codon ACC for residue 43 as Tyr  
 R:Argos, P.; Taylor, W.L.; Minth, C.D.; Dixon, J.E.  
 J. Biol. Chem. 258, 8788-8793, 1983  
 A:Title: Nucleotide and amino acid sequence comparisons of preprosomatostatins.  
 A:Reference number: A01431; MUID:83238516  
 A:Accession: A01431  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <ARG>  
 R:Benoit, R.; Ling, N.; Esch, F.  
 Science 238, 1126-1129, 1987  
 A:Title: A new preprosomatostatin-derived peptide reveals a pattern for prohormone clea

A:Reference number: A01431; MUID:83238516  
 A:Accession: A01431  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <ARG>  
 R:Benoit, R.; Ling, N.; Esch, F.  
 Science 238, 1126-1129, 1987  
 A:Title: A new preprosomatostatin-derived peptide reveals a pattern for prohormone clea

A:Reference number: A47598; MUID:88070564  
 A:Accession: A47598  
 A:Status: Preliminary

A:Molecule type: protein  
 A:Residues: 25-34 <BEN>  
 R:Tavianini, M.A.; Hayes, T.E.; Magazini, M.D.; Minth, C.D.; Dixon, J.E.  
 J. Biol. Chem. 259, 11798-11803, 1984  
 A:Title: Isolation, characterization, and DNA sequence of the rat somatostatin gene.  
 A:Reference number: A22529; MUID:85006903  
 A:Accession: A22529  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-116 <TAV>  
 A:Cross-references: GB:K02248; NID:g207014; PIDN:AAA42161.1; PID:g207017  
 R:Goodman, R.H.; Jacobs, J.W.; Dee, P.C.; Habener, J.F.



J. Biol. Chem. 257, 1156-1159, 1982  
 A:Title: Somatostatin-28 encoded in a cloned cDNA obtained from a rat medullary thyroid  
 A:Reference number: J55220; MUID:82120034  
 A:Accession: J55220  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 38-78; H, 80-116 <RES>  
 A:Cross-references: GB:J00788; NID:g207018; PIDN:AAA42162.1; PID:g207019  
 R:Goodman, R.H.; Montminy, M.R.; Low, M.J.; Habener, J.F.  
 Adv. Exp. Med. Biol. 188, 31-47, 1985  
 A:Title: Biosynthesis of rat preprosomatostatin.  
 A:Reference number: J51829; MUID:85303584  
 A:Accession: J51829  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-116 <RES2>  
 A:Cross-references: GB:M25890; NID:g207030; PIDN:AAA42167.1; PID:g207031  
 C:Comment: Somatostatins are found in a variety of tissues, including hypothalamus, cere  
 C:Genetics: 46/3  
 A:Introns: 46/3  
 A:Function:  
 A:Description: inhibits the secretion of a number of peptide hormones, including somatost  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; hypothalamus; neuropeptide  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-88/Domain: propeptide #status experimental <PRO>  
 F:89-116/Product: somatostatin-28 #status predicted <S28>  
 F:103-116/Product: somatostatin-14 #status predicted <S14>  
 F:105-116/Disulfide bonds: #status predicted

Query Match 61.3%; Score 347; DB 1; Length 116;  
 Best Local Similarity 59.5%; Pred. No. 3.9e-26;  
 Matches 69; Conservative 17; Mismatches 28; Indels 2; Gaps 1;  
 RIPGS  
 somatostatin I precursor - pig (fragment)  
 N:Alternate names: prosomatostatin  
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 30-Nov-1980 #sequence\_revision 31-Jan-1997 #text\_change 31-Jan-1997  
 C:Accession: A34109; A24222; A91273; A93854; A90398; S13616; A01432  
 R:Bersani, M.; Thim, L.; Baldissara, F.G.A.; Holst, J.J.  
 J. Biol. Chem. 264, 10633-10636, 1989  
 A:Title: Prosomatostatin 1-64 is a major product of somatostatin gene expression in f  
 A:Reference number: A34109; MUID:89278131  
 A:Accession: A34109  
 A:Molecule type: protein  
 A:Residues: 1-64 <BER>  
 R:Schmidt, W.E.; Mutt, V.; Kratzin, H.; Carlquist, M.; Conlon, J.M.; Creutzfeldt, W.  
 FEBS Lett. 192, 141-146, 1985  
 A:Title: Isolation and characterization of proSS1-32, a peptide derived from the N-te  
 A:Reference number: A24222; MUID:86030691  
 A:Accession: A24222  
 A:Molecule type: protein  
 A:Residues: 1-32 <SC3>  
 R:Pradayrol, L.; Jornvall, H.; Mutt, V.; Ribet, A.  
 FEBS Lett. 109, 55-58, 1980  
 A:Title: N-terminally extended somatostatin: the primary structure of somatostatin-28  
 A:Reference number: A91273; MUID:80113258  
 A:Accession: A91273  
 A:Molecule type: protein  
 A:Residues: 65-92 <PRO>  
 A:Experimental source: Intestine  
 R:Schally, A.V.; Huang, W.Y.; Chang, R.C.C.; Arimura, A.; Redding, T.W.; Millar, R.P.  
 Proc. Natl. Acad. Sci. U.S.A. 77, 4489-4493, 1980  
 A:Title: Isolation and structure of pro-somatostatin: a putative somatostatin precurs  
 A:Reference number: A93854; MUID:81054799  
 A:Accession: A93854  
 A:Molecule type: protein  
 A:Residues: 65-92 <SCH>  
 R:Schally, A.V.; Dupont, A.; Arimura, A.; Redding, T.W.; Nishi, N.; Linthicum, G.L.;  
 Biochemistry 15, 509-514, 1976  
 A:Title: Isolation and structure of somatostatin from porcine hypothalamus.  
 A:Reference number: A90398; MUID:76136331  
 A:Accession: A90398  
 A:Molecule type: protein  
 A:Residues: 79-92 <SC2>  
 R:Bersani, M.; Johnson, A.H.; Holst, J.J.  
 FEBS Lett. 279, 237-239, 1991  
 A:Title: Oxidation/reduction explains heterogeneity of pancreatic somatostatin.  
 A:Reference number: S13616; MUID:91160722  
 A:Accession: S13616  
 A:Molecule type: protein  
 A:Residues: 79-92 <BE2>  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; hypothalamus; intestine; neuropeptide  
 F:1-64/Domain: propeptide #status experimental <PRO>  
 F:65-92/Product: somatostatin-28 #status experimental <M28>  
 F:79-92/Product: somatostatin-14 #status experimental <M14>  
 F:37/Binding site: carbohydrate (Asn) (covalent) #status absent  
 F:81-92/Disulfide bonds: #status experimental

Query Match 61.3%; Score 347; DB 1; Length 116;  
 Best Local Similarity 59.5%; Pred. No. 3.9e-26;  
 Matches 69; Conservative 17; Mismatches 28; Indels 2; Gaps 1;  
 RIPGS  
 somatostatin precursor - mouse  
 N:Alternate names: preprosomatostatin  
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Jan-1997 #text\_change 18-Jun-1999  
 C:Accession: S08416  
 R:Fuhrmann, G.; Heilig, R.; Kempf, J.; Ebel, A.  
 Nucleic Acids Res. 18, 1287, 1990  
 A:Title: Nucleotide sequence of the mouse preprosomatostatin gene.  
 A:Reference number: S08416; MUID:90206793  
 A:Accession: S08416  
 A:Molecule type: DNA  
 A:Residues: 1-116 <FUH>  
 A:Cross-references: EMBL:X51468; NID:g53768; PIDN:CAA35831.1; PID:g297530  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Genetics: 46/3  
 A:Introns: 46/3  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; neuropeptide  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-88/Domain: propeptide #status predicted <PRO>  
 F:89-116/Product: somatostatin-28 #status predicted <M28>  
 F:103-116/Product: somatostatin-14 #status predicted <M14>  
 F:105-116/Disulfide bonds: #status predicted

Best Local Similarity 59.5%; Pred. No. 3.9e-26;  
 Matches 69; Conservative 17; Mismatches 28; Indels 2; Gaps 1;  
 RIPGS  
 somatostatin I precursor - pig (fragment)  
 N:Alternate names: prosomatostatin  
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 30-Nov-1980 #sequence\_revision 31-Jan-1997 #text\_change 31-Jan-1997  
 C:Accession: A34109; A24222; A91273; A93854; A90398; S13616; A01432  
 R:Bersani, M.; Thim, L.; Baldissara, F.G.A.; Holst, J.J.  
 J. Biol. Chem. 264, 10633-10636, 1989  
 A:Title: Prosomatostatin 1-64 is a major product of somatostatin gene expression in f  
 A:Reference number: A34109; MUID:89278131  
 A:Accession: A34109  
 A:Molecule type: protein  
 A:Residues: 1-64 <BER>  
 R:Schmidt, W.E.; Mutt, V.; Kratzin, H.; Carlquist, M.; Conlon, J.M.; Creutzfeldt, W.  
 FEBS Lett. 192, 141-146, 1985  
 A:Title: Isolation and characterization of proSS1-32, a peptide derived from the N-te  
 A:Reference number: A24222; MUID:86030691  
 A:Accession: A24222  
 A:Molecule type: protein  
 A:Residues: 1-32 <SC3>  
 R:Pradayrol, L.; Jornvall, H.; Mutt, V.; Ribet, A.  
 FEBS Lett. 109, 55-58, 1980  
 A:Title: N-terminally extended somatostatin: the primary structure of somatostatin-28  
 A:Reference number: A91273; MUID:80113258  
 A:Accession: A91273  
 A:Molecule type: protein  
 A:Residues: 65-92 <PRO>  
 A:Experimental source: Intestine  
 R:Schally, A.V.; Huang, W.Y.; Chang, R.C.C.; Arimura, A.; Redding, T.W.; Millar, R.P.  
 Proc. Natl. Acad. Sci. U.S.A. 77, 4489-4493, 1980  
 A:Title: Isolation and structure of pro-somatostatin: a putative somatostatin precurs  
 A:Reference number: A93854; MUID:81054799  
 A:Accession: A93854  
 A:Molecule type: protein  
 A:Residues: 65-92 <SCH>  
 R:Schally, A.V.; Dupont, A.; Arimura, A.; Redding, T.W.; Nishi, N.; Linthicum, G.L.;  
 Biochemistry 15, 509-514, 1976  
 A:Title: Isolation and structure of somatostatin from porcine hypothalamus.  
 A:Reference number: A90398; MUID:76136331  
 A:Accession: A90398  
 A:Molecule type: protein  
 A:Residues: 79-92 <SC2>  
 R:Bersani, M.; Johnson, A.H.; Holst, J.J.  
 FEBS Lett. 279, 237-239, 1991  
 A:Title: Oxidation/reduction explains heterogeneity of pancreatic somatostatin.  
 A:Reference number: S13616; MUID:91160722  
 A:Accession: S13616  
 A:Molecule type: protein  
 A:Residues: 79-92 <BE2>  
 C:Comment: Somatostatin inhibits the release of somatotropin.  
 C:Superfamily: somatostatin  
 C:Keywords: hormone; hypothalamus; intestine; neuropeptide  
 F:1-64/Domain: propeptide #status experimental <PRO>  
 F:65-92/Product: somatostatin-28 #status experimental <M28>  
 F:79-92/Product: somatostatin-14 #status experimental <M14>  
 F:37/Binding site: carbohydrate (Asn) (covalent) #status absent  
 F:81-92/Disulfide bonds: #status experimental

Query Match 51.8%; Score 293; DB 1; Length 92;  
Best Local Similarity 63.0%; Pred. No. 4.3e-21;  
Matches 58; Conservative 13; Mismatches 19; Indels 2; Gaps 1;  
QY 25 APSDAKRLQLQSLMAPAGKQELARNTLVLLSELAHVNEAIEDDMSHGVQEDVDL 84  
DB 1 APSDFRQLQSLAAAGKQELAKYLAELLSEPNQNTDALEPDLUSQAQDEMRL 60  
QY 85 ELENA--PGVPLAPRRKAGCKNFFWKTFTSC 114  
DB 61 ELORSANSNPAPRRKAGCKNFFWKTFTSC 92

RESULT 11  
RIAFSI  
somatostatin I precursor - American goosfish  
N:Species: somatostatin I  
C:Date: 31-Mar-1981 #sequence revision 31-Mar-1981 #text\_change 28-May-1999  
A:Accession: A93236; A93860; A91087; A01433  
R:Hobart, P.; Crawford, R.; Shen, L.; Pictet, R.; Rutter, W.J.  
Nature 288, 137-141, 1980  
A:Title: Cloning and sequence analysis of cDNAs encoding two distinct somatostatin precursor  
A:Reference number: A93236; MUID:81052423  
A:Accession: A93236  
A:Molecule type: mRNA  
A:Residues: 1-121 <HOB>  
A:Cross-references: GB:V00640; GB:J00946; NID:964028; PIDN:CAA23986.1; PID:964029  
R:Goodman, R.H.; Jacobs, J.W.; Chin, W.W.; Lund, P.K.; Dee, P.C.; Habener, J.F.  
Proc. Natl. Acad. Sci. U.S.A. 77, 5869-5873, 1980  
A:Title: Nucleotide sequence of a cloned structural gene coding for a precursor of pancy  
A:Reference number: A93860; MUID:81077276  
A:Accession: A93860  
A:Molecule type: mRNA  
A:Residues: 2-20, 'V', '22-82', 'E', '84-121' <GOO>  
A:Experimental source: islet tissue (endocrine pancreas)  
R:Goodman, R.H.; Jacobs, J.W.; Chin, W.W.; Lund, P.K.; Dee, P.C.; Habener, J.F.  
Proc. Natl. Acad. Sci. U.S.A. 79, 1682, 1982  
A:Reference number: A93905  
A:Contents: annotation; erratum  
R:Noe, B.D.; Spiess, J.; Rivier, J.E.; Vale, W.  
Endocrinology 105, 1410-1415, 1979  
A:Title: Isolation and characterization of somatostatin from anglerfish pancreatic islet  
A:Reference number: A91087; MUID:80046482  
A:Accession: A91087  
A:Molecule type: protein  
A:Residues: 108-121 <NOE>  
C:Superfamily: somatostatin  
C:Keywords: neuropeptide  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:108-121/Product: somatostatin I #status experimental <MAT>  
F:110-121/Disulfide bonds: #status predicted

Query Match 44.3%; Score 251; DB 1; Length 121;  
Best Local Similarity 51.3%; Pred. No. 6.1e-17;  
Matches 60; Conservative 15; Mismatches 36; Indels 6; Gaps 4;

QY 3 STRVQCALA-LLSLALATSSVSAAPSQAKLQLQLQSLMAPAGKQELARNTLVLLSEL 60  
DB 6 SSRRLCLVLLLSLALATSSVSAAPSQAKLQLQLQSLMAPAGKQELARNTLVLLSEL 64  
QY 61 AHVNEAIEDDM--SHGVQEDVDLELERAPGVPLAPRRKAGCKNFFWKTFTSC 114  
DB 65 LOGENEALAEENFPLAEGGPDADLERAAGGGLLAPRRKAGCKNFFWKTFTSC 121

RESULT 12  
151064  
somatostatin II precursor - rainbow trout  
C:Species: Oncorhynchus mykiss (rainbow trout)

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: I51064  
R:Moore, C.A.; Kittilson, J.D.; Dahl, S.K.; Sheridan, M.A.  
Gen Comp Endocrinol 98, 253-261, 1995  
A:Title: Isolation and characterization of a cDNA encoding for preprosomatostatin co  
A:Reference number: I51064; MUID:95354921  
A:Accession: I51064  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-115 <MOO>  
A:Cross-references: EMBL:U32471; NID:9975344; PIDN:AAC59695.1; PID:9975345  
C:Superfamily: somatostatin

Query Match 31.4%; Score 177.5; DB 2; Length 115;  
Best Local Similarity 34.7%; Pred. No. 6e-10;  
Matches 50; Conservative 9; Mismatches 18; Indels 67; Gaps 4;

QY 5 RVOCALALLSLALATSSVSAAPS--DAKLRLQLQLQSLMA----- 41  
DB 5 RIICALLALLGLALATCSOGAASQPDLDLRSRLQLQRAAANPHRSYGVSEWRFTYPNCP 64  
QY 42 -----PAGKQELARNTLVLLSELAHVNEAIEDDMSHGVQEDVDLELERAPGP 92  
DB 65 CLPRKVKCPAG-----AKEDLRVELERSVGN 91  
QY 93 V--LAPRRKAGCKNFFWKTFTSC 114  
DB 92 PNNLPPRRKAGCKNFFWKTFTSC 115

## RESULT 13

RIAFSI  
somatostatin II precursor - American goosfish  
C:Species: Lophius americanus (American goosfish)  
C:Date: 31-Mar-1981 #sequence revision 31-Mar-1981 #text\_change 28-May-1999  
A:Accession: B93236; A94038; A27376; A01434; A21881; A93236  
R:Hobart, P.; Crawford, R.; Shen, L.; Pictet, R.; Rutter, W.J.  
Nature 288, 137-141, 1980  
A:Title: Cloning and sequence analysis of cDNAs encoding two distinct somatostatin p  
A:Reference number: A93236; MUID:81052423  
A:Accession: B93236  
A:Molecule type: mRNA  
A:Residues: 1-125 <HOB>  
A:Cross-references: GB:V00641; GB:J00947; GB:M23199; NID:964030; PIDN:CAA23987.1; PI  
A:Experimental source: islet tissue (endocrine pancreas)  
R:Spies, J.; Noe, B.D.  
Proc. Natl. Acad. Sci. U.S.A. 82, 277-281, 1985  
A:Title: Processing of an anglerfish somatostatin precursor to a hydroxyllysine-conta  
A:Reference number: A94038; MUID:85113184  
A:Accession: A94038  
A:Molecule type: protein  
A:Residues: 98-125 <SHI>  
R:Andrews, P.C.; Nichols, R.; Dixon, J.E.  
J. Biol. Chem. 262, 12692-12699, 1987  
A:Title: Post-translational processing of preprosomatostatin-II examined using fast  
A:Reference number: A27376; MUID:87308304  
A:Accession: A27376  
A:Molecule type: protein  
A:Residues: 1-76, 'DV', '79-89', 'G', '91-125' <AND>  
C:Superfamily: somatostatin  
C:Keywords: hydroxyllysine; neuropeptide; pyroglutamic acid  
F:1-24/Domain: signal sequence #status experimental <SIG>  
F:25-97/Domain: propeptide #status experimental <PRO>  
F:97-125/Product: somatostatin II #status experimental <MAT>  
F:25/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status exper  
F:114-125/Disulfide bonds: #status experimental  
F:120/Modified site: hydroxyllysine (Lys) #status experimental

Query Match 28.7%; Score 162.5; DB 1; Length 125;  
Best Local Similarity 38.9%; Pred. No. 1.8e-08;  
Matches 51; Conservative 16; Mismatches 41; Indels 23; Gaps 6;

QY 1 MLSTVOCALALLSLAATSSVA-----APSDAKLRO--LLQRSIMAPACKOELA 49  
 DB 1 MQCIRCPAILLALLVLCGFSVSSQLDREQSDNQDLDLELRQHWLLERARSAGLLSQEWS 60  
 QY 50 RNTIVELLSEL-----AHVNEAEIEMDMSHGVEQEDVDLEIRAPGPV--LAPRRKAGC 103  
 DB 61 KRAVEELIAQMSLPETFOREA---EDASMATEGR---MNLERSVDSTNNLPPRRKAGC 114  
 QY 104 KNFFWKFTTSC 114  
 DB 115 KNFYWKGFTSC 125

RESULT 14

A61322  
 somatostatin-28 - sheep  
 N:Contains: somatostatin-14  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 07-May-1999  
 Accession: A61322; A61344  
 Spiess, J.; Villarreal, J.; Vale, W.  
 Biochemistry 20, 1982-1988, 1981  
 A:Title: Isolation and sequence analysis of a somatostatin-like polypeptide from ovine  
 A:Reference number: A61322; MUID:81184502  
 A:Accession: A61322  
 A:Molecule type: protein  
 A:Residues: 1-28 <SPI>  
 R:Burgus, R.; Ling, N.; Butcher, M.; Guillemin, R.  
 Proc. Natl. Acad. Sci. U.S.A. 70, 684-688, 1973  
 A:Title: Primary structure of somatostatin, a hypothalamic peptide that inhibits the sec  
 A:Reference number: A61344; MUID:73209562  
 A:Accession: A61344  
 A:Molecule type: protein  
 A:Residues: 15-28 <BUR>  
 C:Superfamily: somatostatin  
 C:Keywords: neuropeptide  
 F:1-28/Product: somatostatin-28 #status experimental <S28>  
 F:15-28/Product: somatostatin-14 #status experimental <SI4>  
 F:17-28/Disulfide bonds: #status experimental

Query Match 22.4%; Score 127; DB 2; Length 28;  
 Best Local Similarity 91.3%; Pred. No. 8e-06; 1; Mismatches 1; Indels 0; Gaps 0;  
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 92 PVLPAPRRKAGCKNFFWKFTTSC 114  
 DB 6 PAMAPRRKAGCKNFFWKFTTSC 28

RESULT 15

A32271  
 somatostatin-related protein - Atlantic hagfish  
 C:Species: Myxine glutinosa (Atlantic hagfish)  
 C:Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 31-Dec-1993  
 Accession: A32271  
 R:Conlon, J.M.; Askensten, U.; Falkmer, S.; Thim, L.  
 Endocrinology 122, 1855-1859, 1988  
 A:Title: Primary structures of somatostatins from the islet organ of the hagfish suggest  
 A:Reference number: A32271; MUID:86195948  
 A:Accession: A32271  
 A:Molecule type: protein  
 A:Residues: 1-34 <CON>  
 C:Superfamily: somatostatin  
 C:Keywords: neuropeptide

Query Match 20.6%; Score 116.5; DB 2; Length 34;  
 Best Local Similarity 60.5%; Pred. No. 0.0001;  
 Matches 23; Conservative 3; Mismatches 7; Indels 5; Gaps 1;

QY 77 VEQEDVDLELERAPGPVLPAPRRKAGCKNFFWKFTTSC 114

DB 2 VERPRQDGVHEPPG-----RERKAGCKNFFWKFTTSC 34

Search completed: June 13, 2002, 12:23:38  
 Job time: 196 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:28 ; Search time 18.73 Seconds  
(without alignments)  
235.666 Million cell updates/sec

Title: US-09-727-739b-3  
Perfect score: 566  
Sequence: 1 MLSTRVQCALALLSLALAISSVSAAPD

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	456	80.6	114	1 SMSL_CARAU	Qygh5 carassius a
2	409	72.3	114	1 SMSL_CARAU	P01171 ictalurus p
3	400	70.7	116	1 SMSL_CARAU	P33094 gallus gall
4	382.5	67.6	115	1 SMSL_CARAU	P87384 ania ridibu
5	373.5	66.0	115	1 SMSL_CARAU	Q9w7f0 protopteris
6	361	63.8	116	1 SMSL_CARAU	P49670 canis fami
7	361	63.8	116	1 SMSL_CARAU	O46688 ovis aries
8	360	63.6	116	1 SMSL_CARAU	P01166 homo sapien
9	352	62.2	116	1 SMSL_CARAU	P26917 bos tauris
10	347	61.3	116	1 SMSL_CARAU	P01167 mus musculu
11	293	51.8	92	1 SMSL_CARAU	P01168 sus scrofa
12	251	44.3	121	1 SMSL_CARAU	P01169 lophius ame
13	177.5	31.4	115	1 SMSL_CARAU	Q91194 oncorhynch
14	163.5	28.9	125	1 SMSL_CARAU	P01170 lophius ame
15	152	27.0	120	1 SMSL_CARAU	Q9ygh4 carassius a
16	129	22.8	126	1 SMSL_CARAU	Q9pr26 ania calva
17	123	21.7	111	1 SMSL_CARAU	Q9ygh3 carassius a
18	117	20.7	109	1 SMSL_CARAU	Q9w7e9 protopteris
19	116.5	20.6	34	1 SMSL_CARAU	P19209 myxine glut
20	114	20.1	73	1 SMSL_CARAU	P21780 platichthys
21	110.5	19.5	105	1 SMSL_CARAU	P01172 ictalurus p
22	108	19.1	35	1 SMSL_CARAU	Q9pr0 lampetra fi
23	107.5	19.0	103	1 SMSL_CARAU	P87385 ania ridibu
24	106	18.7	37	1 SMSL_CARAU	P17779 petromyzon
25	103.5	18.3	74	1 SMSL_CARAU	P09876 myoxocephal
26	99	17.5	28	1 SMSL_CARAU	P81029 oreochromis
27	94.5	16.7	112	1 SMSL_CARAU	O62949 rattus norv
28	87	15.4	14	1 SMSL_CARAU	P20750 myoxocephal
29	87	15.4	14	1 SMSL_CARAU	P31865 alligator m
30	86.5	15.3	105	1 SMSL_CARAU	O00230 homo sapien
31	85	15.0	109	1 SMSL_CARAU	P36469 mus musculu
32	74	13.1	297	1 HTPX_STRGC	O30795 streptococc
33	73	12.9	493	1 VPE_VICSA	P49044 vicia sativ

34	72	12.7	256	1	CFL_ELAUM	O65333 elaeagnus u
35	72	12.7	1436	1	MRPS_RAT	Q9gym0 rattus norv
36	70.5	12.5	1087	1	AKA9_RABIT	Q28628 myctolagus
37	70.5	12.5	3067	1	CAIC_MOUSE	Q08047 mus musculu
38	69.5	12.3	616	1	MUTA_STRCM	Q05064 streptomyce
39	69	12.2	98	1	SNC5_HUMAN	O75971 homo sapien
40	69	12.2	402	1	OPS4_CANAL	P46596 candida alb
41	69	12.2	453	1	SSF2_YEAST	Q12153 saccharomyc
42	69	12.2	646	1	NODQ_RHIS3	P72339 r nodq bifu
43	68.5	12.1	442	1	MFAL_CHICK	P35080 gallus gall
44	68	12.0	327	1	FRA2_RAT	P51145 rattus norv
45	68	12.0	396	1	Y858_TREPA	O83830 treponema p

## ALIGNMENTS

RESULT 1  
SMSA\_CARAU STANDARD; PRT; 114 AA.  
AC QYGH5;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Somatostatin IA precursor [Contains: Somatostatin-26; Somatostatin-14]  
DE Carassius auratus (Goldfish).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Carassius.  
OC NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Lin X.-W., Peter R.E.;  
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I and -II from goldfish brain."  
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: U40754; AAD09359.1;  
DR InterPro: IPR004250; Somatostatin.  
DR Pfam: PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
FT SIGNAL  
FT PROPEP 25 88  
FT PEPTIDE 89 114  
FT PEPTIDE 101 114  
FT DISULFID 103 114  
FT BY SIMILARITY  
SQ SEQUENCE 114 AA; 12574 MW; B5920015E2D272A4 CRC64;

Query Match 80.6%; Score 456; DB 1; Length 114;  
Best Local Similarity 78.9%; Pred. No. 2,1e-36;  
Matches 90; Conservative 12; Mismatches 12; Indels 0; Gaps 0;  
Qy 1 MLSTRVQCALALLSLALAISSVSAAPDVKRLQLLRSMLAPAGQELARTLVLLSEL 60  
Dy 1 MLSTRVQCALALLSLALAISSVSAAPDVKRLQLLRSMLAPAGQELARTLVLLSEL 60  
Qy 61 ARHVEAIEIDMSHGVEQEDVLELAPGVPVLAPEPRKAGCKNFFWKTFTSC 114  
Dy 61 VQANEALPEDLSRAVEKDEVRLELAPGVPVLAPEPRKAGCKNFFWKTFTSC 114



DE 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin I precursor (PSS1) [Contains: Somatostatin-14 (S-I)  
(SSS1)]  
OS Rana ridibunda (laughing frog) (Marsh frog).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoides; Ranidae; Rana.  
OX NCBI\_TaxID=8406;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=97057290; PubMed=8901629;  
RA Testivint H., Lihmann I., Buchary C., Vieau D., Coulouarn Y.,  
RA Fournier A., Conlon J.M., Vaudry H.;  
RT "Occurrence of two somatostatin variants in the frog brain:  
RT characterization of the cDNAs, distribution of the mRNAs, and  
RT receptor-binding affinities of the peptides".  
RL Proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).  
[2]  
[3]  
[4]  
[5]  
[6]  
[7]  
[8]  
[9]  
[10]  
[11]  
[12]  
[13]  
[14]  
[15]  
[16]  
[17]  
[18]  
[19]  
[20]  
[21]  
[22]  
[23]  
[24]  
[25]  
[26]  
[27]  
[28]  
[29]  
[30]  
[31]  
[32]  
[33]  
[34]  
[35]  
[36]  
[37]  
[38]  
[39]  
[40]  
[41]  
[42]  
[43]  
[44]  
[45]  
[46]  
[47]  
[48]  
[49]  
[50]  
[51]  
[52]  
[53]  
[54]  
[55]  
[56]  
[57]  
[58]  
[59]  
[60]  
[61]  
[62]  
[63]  
[64]  
[65]  
[66]  
[67]  
[68]  
[69]  
[70]  
[71]  
[72]  
[73]  
[74]  
[75]  
[76]  
[77]  
[78]  
[79]  
[80]  
[81]  
[82]  
[83]  
[84]  
[85]  
[86]  
[87]  
[88]  
[89]  
[90]  
[91]  
[92]  
[93]  
[94]  
[95]  
[96]  
[97]  
[98]  
[99]  
[100]  
[101]  
[102]  
[103]  
[104]  
[105]  
[106]  
[107]  
[108]  
[109]  
[110]  
[111]  
[112]  
[113]  
[114]  
[115]  
[116]  
[117]  
[118]  
[119]  
[120]  
[121]  
[122]  
[123]  
[124]  
[125]  
[126]  
[127]  
[128]  
[129]  
[130]  
[131]  
[132]  
[133]  
[134]  
[135]  
[136]  
[137]  
[138]  
[139]  
[140]  
[141]  
[142]  
[143]  
[144]  
[145]  
[146]  
[147]  
[148]  
[149]  
[150]  
[151]  
[152]  
[153]  
[154]  
[155]  
[156]  
[157]  
[158]  
[159]  
[160]  
[161]  
[162]  
[163]  
[164]  
[165]  
[166]  
[167]  
[168]  
[169]  
[170]  
[171]  
[172]  
[173]  
[174]  
[175]  
[176]  
[177]  
[178]  
[179]  
[180]  
[181]  
[182]  
[183]  
[184]  
[185]  
[186]  
[187]  
[188]  
[189]  
[190]  
[191]  
[192]  
[193]  
[194]  
[195]  
[196]  
[197]  
[198]  
[199]  
[200]  
[201]  
[202]  
[203]  
[204]  
[205]  
[206]  
[207]  
[208]  
[209]  
[210]  
[211]  
[212]  
[213]  
[214]  
[215]  
[216]  
[217]  
[218]  
[219]  
[220]  
[221]  
[222]  
[223]  
[224]  
[225]  
[226]  
[227]  
[228]  
[229]  
[230]  
[231]  
[232]  
[233]  
[234]  
[235]  
[236]  
[237]  
[238]  
[239]  
[240]  
[241]  
[242]  
[243]  
[244]  
[245]  
[246]  
[247]  
[248]  
[249]  
[250]  
[251]  
[252]  
[253]  
[254]  
[255]  
[256]  
[257]  
[258]  
[259]  
[260]  
[261]  
[262]  
[263]  
[264]  
[265]  
[266]  
[267]  
[268]  
[269]  
[270]  
[271]  
[272]  
[273]  
[274]  
[275]  
[276]  
[277]  
[278]  
[279]  
[280]  
[281]  
[282]  
[283]  
[284]  
[285]  
[286]  
[287]  
[288]  
[289]  
[290]  
[291]  
[292]  
[293]  
[294]  
[295]  
[296]  
[297]  
[298]  
[299]  
[300]  
[301]  
[302]  
[303]  
[304]  
[305]  
[306]  
[307]  
[308]  
[309]  
[310]  
[311]  
[312]  
[313]  
[314]  
[315]  
[316]  
[317]  
[318]  
[319]  
[320]  
[321]  
[322]  
[323]  
[324]  
[325]  
[326]  
[327]  
[328]  
[329]  
[330]  
[331]  
[332]  
[333]  
[334]  
[335]  
[336]  
[337]  
[338]  
[339]  
[340]  
[341]  
[342]  
[343]  
[344]  
[345]  
[346]  
[347]  
[348]  
[349]  
[350]  
[351]  
[352]  
[353]  
[354]  
[355]  
[356]  
[357]  
[358]  
[359]  
[360]  
[361]  
[362]  
[363]  
[364]  
[365]  
[366]  
[367]  
[368]  
[369]  
[370]  
[371]  
[372]  
[373]  
[374]  
[375]  
[376]  
[377]  
[378]  
[379]  
[380]  
[381]  
[382]  
[383]  
[384]  
[385]  
[386]  
[387]  
[388]  
[389]  
[390]  
[391]  
[392]  
[393]  
[394]  
[395]  
[396]  
[397]  
[398]  
[399]  
[400]  
[401]  
[402]  
[403]  
[404]  
[405]  
[406]  
[407]  
[408]  
[409]  
[410]  
[411]  
[412]  
[413]  
[414]  
[415]  
[416]  
[417]  
[418]  
[419]  
[420]  
[421]  
[422]  
[423]  
[424]  
[425]  
[426]  
[427]  
[428]  
[429]  
[430]  
[431]  
[432]  
[433]  
[434]  
[435]  
[436]  
[437]  
[438]  
[439]  
[440]  
[441]  
[442]  
[443]  
[444]  
[445]  
[446]  
[447]  
[448]  
[449]  
[450]  
[451]  
[452]  
[453]  
[454]  
[455]  
[456]  
[457]  
[458]  
[459]  
[460]  
[461]  
[462]  
[463]  
[464]  
[465]  
[466]  
[467]  
[468]  
[469]  
[470]  
[471]  
[472]  
[473]  
[474]  
[475]  
[476]  
[477]  
[478]  
[479]  
[480]  
[481]  
[482]  
[483]  
[484]  
[485]  
[486]  
[487]  
[488]  
[489]  
[490]  
[491]  
[492]  
[493]  
[494]  
[495]  
[496]  
[497]  
[498]  
[499]  
[500]  
[501]  
[502]  
[503]  
[504]  
[505]  
[506]  
[507]  
[508]  
[509]  
[510]  
[511]  
[512]  
[513]  
[514]  
[515]  
[516]  
[517]  
[518]  
[519]  
[520]  
[521]  
[522]  
[523]  
[524]  
[525]  
[526]  
[527]  
[528]  
[529]  
[530]  
[531]  
[532]  
[533]  
[534]  
[535]  
[536]  
[537]  
[538]  
[539]  
[540]  
[541]  
[542]  
[543]  
[544]  
[545]  
[546]  
[547]  
[548]  
[549]  
[550]  
[551]  
[552]  
[553]  
[554]  
[555]  
[556]  
[557]  
[558]  
[559]  
[560]  
[561]  
[562]  
[563]  
[564]  
[565]  
[566]  
[567]  
[568]  
[569]  
[570]  
[571]  
[572]  
[573]  
[574]  
[575]  
[576]  
[577]  
[578]  
[579]  
[580]  
[581]  
[582]  
[583]  
[584]  
[585]  
[586]  
[587]  
[588]  
[589]  
[590]  
[591]  
[592]  
[593]  
[594]  
[595]  
[596]  
[597]  
[598]  
[599]  
[600]  
[601]  
[602]  
[603]  
[604]  
[605]  
[606]  
[607]  
[608]  
[609]  
[610]  
[611]  
[612]  
[613]  
[614]  
[615]  
[616]  
[617]  
[618]  
[619]  
[620]  
[621]  
[622]  
[623]  
[624]  
[625]  
[626]  
[627]  
[628]  
[629]  
[630]  
[631]  
[632]  
[633]  
[634]  
[635]  
[636]  
[637]  
[638]  
[639]  
[640]  
[641]  
[642]  
[643]  
[644]  
[645]  
[646]  
[647]  
[648]  
[649]  
[650]  
[651]  
[652]  
[653]  
[654]  
[655]  
[656]  
[657]  
[658]  
[659]  
[660]  
[661]  
[662]  
[663]  
[664]  
[665]  
[666]  
[667]  
[668]  
[669]  
[670]  
[671]  
[672]  
[673]  
[674]  
[675]  
[676]  
[677]  
[678]  
[679]  
[680]  
[681]  
[682]  
[683]  
[684]  
[685]  
[686]  
[687]  
[688]  
[689]  
[690]  
[691]  
[692]  
[693]  
[694]  
[695]  
[696]  
[697]  
[698]  
[699]  
[700]  
[701]  
[702]  
[703]  
[704]  
[705]  
[706]  
[707]  
[708]  
[709]  
[710]  
[711]  
[712]  
[713]  
[714]  
[715]  
[716]  
[717]  
[718]  
[719]  
[720]  
[721]  
[722]  
[723]  
[724]  
[725]  
[726]  
[727]  
[728]  
[729]  
[730]  
[731]  
[732]  
[733]  
[734]  
[735]  
[736]  
[737]  
[738]  
[739]  
[740]  
[741]  
[742]  
[743]  
[744]  
[745]  
[746]  
[747]  
[748]  
[749]  
[750]  
[751]  
[752]  
[753]  
[754]  
[755]  
[756]  
[757]  
[758]  
[759]  
[760]  
[761]  
[762]  
[763]  
[764]  
[765]  
[766]  
[767]  
[768]  
[769]  
[770]  
[771]  
[772]  
[773]  
[774]  
[775]  
[776]  
[777]  
[778]  
[779]  
[780]  
[781]  
[782]  
[783]  
[784]  
[785]  
[786]  
[787]  
[788]  
[789]  
[790]  
[791]  
[792]  
[793]  
[794]  
[795]  
[796]  
[797]  
[798]  
[799]  
[800]  
[801]  
[802]  
[803]  
[804]  
[805]  
[806]  
[807]  
[808]  
[809]  
[810]  
[811]  
[812]  
[813]  
[814]  
[815]  
[816]  
[817]  
[818]  
[819]  
[820]  
[821]  
[822]  
[823]  
[824]  
[825]  
[826]  
[827]  
[828]  
[829]  
[830]  
[831]  
[832]  
[833]  
[834]  
[835]  
[836]  
[837]  
[838]  
[839]  
[840]  
[841]  
[842]  
[843]  
[844]  
[845]  
[846]  
[847]  
[848]  
[849]  
[850]  
[851]  
[852]  
[853]  
[854]  
[855]  
[856]  
[857]  
[858]  
[859]  
[860]  
[861]  
[862]  
[863]  
[864]  
[865]  
[866]  
[867]  
[868]  
[869]  
[870]  
[871]  
[872]  
[873]  
[874]  
[875]  
[876]  
[877]  
[878]  
[879]  
[880]  
[881]  
[882]  
[883]  
[884]  
[885]  
[886]  
[887]  
[888]  
[889]  
[890]  
[891]  
[892]  
[893]  
[894]  
[895]  
[896]  
[897]  
[898]  
[899]  
[900]  
[901]  
[902]  
[903]  
[904]  
[905]  
[906]  
[907]  
[908]  
[909]  
[910]  
[911]  
[912]  
[913]  
[914]  
[915]  
[916]  
[917]  
[918]  
[919]  
[920]  
[921]  
[922]  
[923]  
[924]  
[925]  
[926]  
[927]  
[928]  
[929]  
[930]  
[931]  
[932]  
[933]  
[934]  
[935]  
[936]  
[937]  
[938]  
[939]  
[940]  
[941]  
[942]  
[943]  
[944]  
[945]  
[946]  
[947]  
[948]  
[949]  
[950]  
[951]  
[952]  
[953]  
[954]  
[955]  
[956]  
[957]  
[958]  
[959]  
[960]  
[961]  
[962]  
[963]  
[964]  
[965]  
[966]  
[967]  
[968]  
[969]  
[970]  
[971]  
[972]  
[973]  
[974]  
[975]  
[976]  
[977]  
[978]  
[979]  
[980]  
[981]  
[982]  
[983]  
[984]  
[985]  
[986]  
[987]  
[988]  
[989]  
[990]  
[991]  
[992]  
[993]  
[994]  
[995]  
[996]  
[997]  
[998]  
[999]  
[1000]

Somatostatin-14].  
DE Prototestis-annectens (African lungfish).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Dipnoi; Lepidosireniaformes; Protopterygia; Protopterus.  
OX NCBI\_TaxID=7888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99326690; PubMed=10398054;  
RA Trabucchi M., Testivint H., Lihmann I., Jegou S., Vallarino M.,  
RA Vaudry H.;  
RT "Molecular cloning of the cDNAs and distribution of the mRNAs encoding  
RT two somatostatin precursors in the African lungfish *Protopterus*  
RT *annectens*".  
RL J. Comp. Neurol. 410:643-652(1999).  
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF126243; AAD39138.1;  
DR InterPro: IPR004250; Somatostatin.  
DR Pfam: PF03002; Somatostatin; 1.  
KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT PROPEP 25 88  
FT PEPTIDE 89 115 SOMATOSTATIN-27 (POTENTIAL).  
FT PEPTIDE 102 115 SOMATOSTATIN-14.  
FT DISULFID 104 115 BY SIMILARITY.  
FT SEQUENCE 115 AA; 12600 MW; B0CEFE603FEAF09 CRC64;  
SQ  
Query Match 66.08; Score 373.5; DB 1; Length 115;  
Best Local Similarity 67.88; Pred. No. 1.2e-28;  
Matches 78; Conservative 12; Mismatches 24; Indels 1; Gaps 1;  
QY 1 MLSTRVOCALALLSLAIAISSVSAAPSDAKLQRLQSLMAPAGKQELARNTLVLLSEL 60  
Db 1 MLSCRFQCALVLLSLAVVFSKVSAPSDLRQLQSLAAGKQELKYSLLSEL 60  
QY 61 AHVNEAIELDMSHGVEQEDVLELER--APGVLAPRRKAGCKNFFWKTFTSC 114  
Db 61 AQSENDALDSDLSRGADQDEVRLDRLSANSPLAARRKAGCKNFFWKTFTSC 115  
RESULT 6  
SMS\_CANFA STANDARD; PRT; 116 AA.  
AC P49670;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
GN SST.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Gastric mucosa;  
RX MEDLINE=97142297; PubMed=8988514;  
RA Dickinson C.J., Delvalle J., Rodisco A., Gantz I., Tong L.,  
RA Finniss S., Yamada T.;  
RT "Canine prosomatostatin: isolation of a cDNA, regulation of gene  
RT expression, and characterization of post-translational processing





CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC EMBL: J00306; AAA60566.1;  
 CC EMBL: M19318; AAA36908.1;  
 CC PIR: A43614; RHUS1;  
 CC PIR: A38968; A28968;  
 CC MIM: 182450;  
 CC InterPro: IPR004250; Somatostatin.  
 CC Pfam: PF03002; Somatostatin; 1.  
 CC Cleavage on pair of basic residues; Hormone; Signal; Pharmaceutical.  
 FT SIGNAL 1 24  
 FT PROPEP 25 88  
 FT PEPTIDE 89 116 SOMATOSTATIN-28.  
 FT PEPTIDE 103 116 SOMATOSTATIN-14.  
 FT DISULFID 105 116  
 FT SEQUENCE 116 AA; 12735 MW; AB49BB89DC9DD8DA CRC64;

Query Match 63.6%; Score 360; DB 1; Length 116;  
 Best Local Similarity 62.1%; Pred. No. 2.3e-27;  
 Matches 72; Conservative 17; Mismatches 25; Indels 2; Gaps 1;  
 QY 1 MLSTVQCALALLSLAIAISSVSAAPSDAKLRLQRLSRLMAPAGKQELARNTLVLLSEL 60  
 DB 1 MLSCRLQCALAALSIVLALGCVTGAPSDPLRQFLQKSLAAAGKQELAKYFLAELLSEP 60  
 QY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PCPVLAPEPRKAGCKNFFWKTFTSC 114  
 DB 61 NOTENDALEPELDSQAAEODEMRLELQRSANSNPAMPAPRRKAGCKNFFWKTFTSC 116

RESULT 9  
 SMS\_BOVIN  
 ID SMS\_BOVIN STANDARD; PRT; 116 AA.  
 AC P26917;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].  
 GN SST OR SMST.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88288237; PubMed=2899837;  
 RA Su C.J., White J.W., Li W.H., Luo C.C., Frazier M.L., Saunders G.F.,  
 Chan L.;  
 RT "Structure and evolution of somatostatin genes.";  
 RL Mol. Endocrinol. 2:209-216(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HOLSTEIN;  
 RX Furu L.M., Kazmer G.W., Strausbaugh L., Zinn S.A.;  
 RT "Cloning and characterization of the bovine somatostatin gene.";  
 RL J. Anim. Sci. 77:492-493(1993).  
 CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC EMBL: M31217; AAA30744.1;  
 CC EMBL: U97077; AAB58056.1;  
 CC PIR: A40929; A40929;  
 CC InterPro: IPR004250; Somatostatin.  
 CC Pfam: PF03002; Somatostatin; 1.  
 CC Cleavage on pair of basic residues; Hormone; Signal.  
 FT SIGNAL 1 24  
 FT PROPEP 25 88  
 FT PEPTIDE 89 116 SOMATOSTATIN-28.  
 FT PEPTIDE 103 116 SOMATOSTATIN-14.  
 FT DISULFID 105 116  
 FT SEQUENCE 116 AA; 12688 MW; C18F17E64A371D8E CRC64;

Query Match 62.2%; Score 352; DB 1; Length 116;  
 Best Local Similarity 61.2%; Pred. No. 1.3e-26;  
 Matches 71; Conservative 17; Mismatches 26; Indels 2; Gaps 1;  
 QY 1 MLSTVQCALALLSLAIAISSVSAAPSDAKLRLQRLSRLMAPAGKQELARNTLVLLSEL 60  
 DB 1 MLSCRLQCALAALSIVLALGCVTGAPSDPLRQFLQKSLAAAGKQELAKYFLAELLSEP 60  
 QY 61 AHVENEATELDDMSHGVEQEDVDLELERA--PCPVLAPEPRKAGCKNFFWKTFTSC 114  
 DB 61 NOTENDALEPELDSQAAEODEMRLELQRSANSNPAMPAPRRKAGCKNFFWKTFTSC 116

RESULT 10  
 SMS\_MOUSE  
 ID SMS\_MOUSE STANDARD; PRT; 116 AA.  
 AC P01167;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Somatostatin precursor [Contains: Antrin; Somatostatin-28;  
 DE Somatostatin-14].  
 GN SST OR SMST.  
 OS Mus musculus (Mouse), and  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090, 10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX SPECIES=Rat;  
 RX MEDLINE=85006903; PubMed=6148343;  
 RA Tavianini M.A., Hayes T.E., Magazin M.D., Minth C.D., Dixon J.E.;  
 RT "Isolation, characterization, and DNA sequence of the rat  
 RT somatostatin gene.";  
 RL J. Biol. Chem. 259:11798-11803(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Rat;  
 RX MEDLINE=85238516; PubMed=6134734;  
 RA Argos P., Taylor W.L., Minth C.D., Dixon J.E.;  
 RT "Nucleotide and amino acid sequence comparisons of  
 RT preprosomatostatin.";  
 RL J. Biol. Chem. 258:8788-8793(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Rat;  
 RX MEDLINE=83213516; PubMed=6133871;  
 RA Goodman R.H., Aron D.C., Roos B.A.;  
 RT "Rat pre-prosomatostatin. Structure and processing by microsomal  
 RT membranes.";  
 RL J. Biol. Chem. 258:5570-5573(1983).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Rat;  
 RX MEDLINE=85303584; PubMed=2863939;  
 RA Goodman R.H., Montminy M.R., Low M.J., Habener J.F.;



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U36385; AAB38485.1;  
 CC PIR: A01433; RIPS.  
 CC PIR: A24222; A24222.  
 CC PIR: A34109; A34109.  
 CC PIR: S13616; S13616.  
 CC InterPro: IPR004250; Somatostatin.  
 CC Pfam: PF03002; Somatostatin; 1.  
 CC KW Cleavage on pair of basic residues; Hormone.  
 CC NON\_TER 1 1  
 CC FT PROPEP 1 64  
 CC FT PEPTIDE 65 92 SOMATOSTATIN-28.  
 CC FT PEPTIDE 79 92 SOMATOSTATIN-14.  
 CC FT DISULFID 81 92  
 CC SEQUENCE 92 AA: 10346 MW: 787CBE82CFBBAE76 CRC64;

Query Match 51.8%; Score 293; DB 1; Length 92;  
 Best Local Similarity 63.0%; Pred. No. 3.7e-21;  
 Matches 58; Conservative 13; Mismatches 19; Indels 2; Gaps 1;  
 Qy 25 APSDAKLROLLRSAPAGKQELARNLTVLLSELAHVNEAIEDDMHGVQEDVDL 84  
 Db 1 APSDRLRFLOKSLAAAGKQELAKYFLAELLSEPNQENDALEPEDLQAABQDEMRL 60  
 Qy 85 ELERA--PGPVLAPEPRKAGCKNFFWKTFTSC 114  
 Db 61 ELORSANSNPAMPREPRKAGCKNFFWKTFTSC 92

RESULT 12  
 ID SMSL\_LOPAM STANDARD; PRT; 121 AA.  
 AC P01169;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Somatostatin I precursor [Contains: Somatostatin-14].  
 OS Lophius americanus (American goosefish) (Anglerfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.  
 NCBI\_TaxID=8073;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 X MEDLINE=81052423; PubMed=6107860;  
 A Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;  
 RT "Cloning and sequence analysis of cDNAs encoding two distinct  
 RT somatostatin precursors found in the endocrine pancreas of  
 RT anglerfish";  
 RT Nature 288:137-141(1980).  
 RL [2]  
 RN SEQUENCE OF 2-121 FROM N.A.  
 RP MEDLINE=8107276; PubMed=6108560;  
 RA Goodman R.H., Jacobs J.W., Chin W.W., Lund P.K., Dee P.C.,  
 RA Habener J.F.;  
 RT "Nucleotide sequence of a cloned structural gene coding for a  
 RT precursor of pancreatic somatostatin";  
 RL Proc. Natl. Acad. Sci. U.S.A. 77:5869-5873(1980).  
 RN [3]  
 RP ERRATUM.  
 RA Goodman R.H., Jacobs J.W., Chin W.W., Lund P.K., Dee P.C.,  
 RA Habener J.F.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1682-1682(1982).  
 RN [4]  
 RP SEQUENCE OF 108-121.  
 RX MEDLINE=80046482; PubMed=387385;  
 RA Noe B.D., Spiess J., Rivier J.E., Vale W.;  
 RT "Isolation and characterization of somatostatin from anglerfish  
 RT pancreatic islet";  
 RL Endocrinology 105:1410-1415(1979).

CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: V00640; CAA23986.1;  
 CC PIR: A01433; RIAFSI.  
 CC InterPro: IPR004250; Somatostatin.  
 CC Pfam: PF03002; Somatostatin; 1.  
 CC KW Cleavage on pair of basic residues; Hormone; Signal;  
 CC Multigene family.  
 CC SIGNAL 1 24 PROBABLE.  
 CC FT PROPEP 25 105  
 CC FT PEPTIDE 108 121 SOMATOSTATIN-14.  
 CC FT DISULFID 110 121  
 CC FT CONFLICT 21 21 A -> V (IN REF. 2).  
 CC FT CONFLICT 83 83 G -> E (IN REF. 2).  
 CC FT CONFLICT 83 83  
 CC SEQUENCE 121 AA: 13325 MW: D70C53DC798C2095 CRC64;

Query Match 44.3%; Score 251; DB 1; Length 121;  
 Best Local Similarity 51.3%; Pred. No. 4.5e-17;  
 Matches 60; Conservative 15; Mismatches 36; Indels 6; Gaps 4;

Qy 3 STVOCALAL-LLSLALATSSVSAAPSDAKLROLLRSAPAGKQELARNLTVLLSEL 60  
 Db 6 SSRRLCVLLLSLTASISCSFAGQDSKRLRLHRYPL-QGSKDMTRSLAELLSDL 64  
 Qy 61 AHVNEAIEIDDM---SHGVQEDVDLELERAGPVLAPRERKAGCKNFFWKTFTSC 114  
 Db 65 LOGENEALIEENFPLAEGGPEDAHADLERAAAGGPLLAPRERKAGCKNFFWKTFTSC 121

RESULT 13  
 SMS2\_ONCMY  
 ID SMS2\_ONCMY STANDARD; PRT; 115 AA.  
 AC Q91194;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;  
 DE [Tyr7,Gly10]somatostatin-14].  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95354921; PubMed=7628684;  
 RA Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;  
 RT "Isolation and characterization of a cDNA encoding for  
 RT preprosomatostatin containing [Tyr7, Gly10]-somatostatin-14 from the  
 RT endocrine pancreas of rainbow trout, *Oncorhynchus mykiss*.";  
 RL Gen. Comp. Endocrinol. 98:253-261(1995).  
 CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; U32471; AAC59695.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 87
FT PEPTIDE 88 115 [TYR21, GLY24]SOMATOSTATIN-28 (POTENTIAL).
FT DISULFID 102 115 [TYR7, GLY10]SOMATOSTATIN-14.
FT SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;
SQ

Query Match 31.4%; Score 177.5; DB 1; Length 115;
Best Local Similarity 34.7%; Pred No. 3.6e-10;
Matches 50; Conservative 9; Mismatches 18; Indels 67; Gaps 4;

QY 5 RVQCALALLSLALAISSVSAAPS---DAKLRLQLRSLMA----- 41
DB 5 RIHCALALLGLALICSGAASQPDLLRGRLLRLQRAAAWPHRSGYSERWRTFFPNCP 64
42 -----PAGKQELARTLVLLSELARHVEAEIELDDMSHGVEQEDVDLELRAPGP 92
65 CLRPRKVKCPAG-----AKEDLRVELERSVGN 91
QY 93 V--LAPRRKAGCKNFKWTFSC 114
DB 92 PNNLPFRKAGCKNFKWTFSC 115

RESULT 14
SMS2_LOPAM STANDARD; PRT; 125 AA.
AC P01170; Q91066;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr7, Gly10]somatostatin-14].
OS Lophius americanus (American goosefish) (Anglerfish).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX NCBI_TaxID=8073;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=61052423; PubMed=6107860;
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;
FT "Cloning and sequence analysis of cDNAs encoding two distinct somatostatin precursors found in the endocrine pancreas of anglerfish."
AL Nature 288:137-141(1990).
RN [2]
RP PARTIAL SEQUENCE, AND HYDROXYLATION.
RX MEDLINE=87308304; PubMed=2887572;
RA Andrews P.C., Nichols R., Dixon J.E.;
RT "Post-translational processing of preprosomatostatin-II examined using fast atom bombardment mass spectrometry."
RL J. Biol. Chem. 262:12692-12699(1987).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; V00641; CAA23987.1;

```

```

DR PIR; A01434; RIAFS2.
DR PIR; A27376; A27376.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 109 [TYR7, GLY10]SOMATOSTATIN-14.
FT PEPTIDE 112 125
FT DISULFID 114 125 HYDROXYLATION.
FT MOD_RES 120 120 DV -> TG (IN REF. 1).
FT CONFLICT 77 78
FT CONFLICT 90 90 G -> E (IN REF. 1).
SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 28.9%; Score 163.5; DB 1; Length 125;
Best Local Similarity 38.2%; Pred No. 8.2e-09;
Matches 50; Conservative 17; Mismatches 41; Indels 23; Gaps 5;

QY 1 MLSTRVOCALALLSLALAISSVSA-----APSDAKLRQ--LLQKSLMAPAGKQELA 49
DB 1 MQCIRCPAILALLALVLCGPSVSQDREQSDNQDLDLELRHLLERARSAGLLSOWS 60
QY 50 RNTLVLLSEL-----AHVENEAEIELDDMSHGVEQEDVDLELRAPGPV--LAPRRKAGC 103
DB 61 KRAVEELLAQMSLPADVQREAEADSMATGG-----RMNLSRVSVDSTNNLPFRKAGC 114
QY 104 KNFFKWTFTSC 114
DB 115 KNFYKGTFTSC 125

RESULT 15
SMS2_CARAU STANDARD; PRT; 120 AA.
ID SMS2_CARAU
AC Q9YGH4; Q9PT02;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21, Gly24]somatostatin-28].
DE [Tyr7, Gly10]somatostatin-14].
DE Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I and -II from goldfish brain."
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Otto C.J., Peter R.E.;
RT "The expression of SRIF mRNA in the brain of goldfish."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; U60262; AAD09626.1;

```

DR EMBL: AF025686; AAF15306.1; -  
 DR InterPro: IPR004250; Somatostatin.  
 DR Pfam: PF03002; Somatostatin; 1.  
 KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 92  
 FT PEPTIDE 93 120 [TYR21, GLY24] SOMATOSTATIN-28.  
 FT PEPTIDE 107 120 [TYR7, GLY10] SOMATOSTATIN-14.  
 FT DISULFID 109 120 BY SIMILARITY.  
 FT CONFLICT 51 51 Q -> RW (IN REF. 2).  
 SQ SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;

Query Match 27.0%; Score 153; DB 1; Length 120;  
 Best Local Similarity 37.9%; Pred. No. 7.6e-08;  
 Matches 44; Conservative 16; Mismatches 48; Indels 8; Gaps 3;  
 QY 6 VOCALALLSLALAI-----SSVSAAPSDAKLROLLQRLSILMAPAGKQELARNLYVELLSEL 60  
 b 6 LHCYLALLGLSLVLCGRCANSQLEPDLDLFRHRLQLQASATGQATQDFTKRDVEKLLSL 65  
 QY 61 AHVNEAIELDDMSHGVEQEDVDLELERA--PGVLAPRRKAGCKNFFWTFYTC 114  
 Db 66 SIFEMEMRE-KGLSMAGESEDLRLQERSAESNNQLPTRVRKEGCKNFYWKGTSC 120

Search completed: June 13, 2002, 12:29:53  
 Job time: 445 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 12:22:58 ; Search time 60.72 seconds  
(without alignments)  
324,793 Million cell updates/sec

Title: US-09-727-739b-3  
Perfect score: 566  
Sequence: 1 MLSTRVQCALALLSLALAISSVSAAPSADKRLQRLQSLRMAPAGKQELARNTLVLLSEL 60

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SPTREMBL19:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_rv:  
16: sp\_bacteriophage:  
17: sp\_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	456	80.6	114	13 Q90Y41	Q90Y41 gnathoniemus
2	454	80.2	114	13 Q90Y40	Q90Y40 chitalla chi
3	400	70.7	116	13 Q90XEL	Q90XEL acipenser t
4	208	36.7	116	13 Q90Y42	Q90Y42 pantodon bu
5	201.5	35.6	115	13 Q90Y43	Q90Y43 osteoglossu
6	196	34.6	120	13 Q90Y39	Q90Y39 catostomus
7	117	20.7	111	13 Q90XEO	Q90XEO acipenser t
8	107.5	19.0	107	13 Q90XEO	Q90XEO brachydanio
9	99	17.5	28	13 Q90PRN9	Q90PRN9 carassius a
10	87	15.4	25	13 Q90PRV0	Q90PRV0 anquilla ja
11	85	15.0	105	11 Q90L1P8	Q90L1P8 mus musculus
12	79.5	14.0	715	11 Q63803	Q63803 rattus norv
13	77	13.6	376	5 Q01392	Q01392 caenorhabdi
14	77	13.6	385	5 Q17355	Q17355 caenorhabdi
15	75	13.3	323	5 Q98345	Q98345 leishmania
16	74.5	13.2	422	11 Q90JX0	Q90JX0 mus musculus

17	74.5	13.2	756	11 Q921N8	Q921N8 mus musculu
18	74	13.1	557	5 Q961H0	Q961H0 drosophila
19	74	13.1	1385	5 Q9V8Z5	Q9V8Z5 drosophila
20	74	13.1	1389	5 Q24591	Q24591 drosophila
21	74	13.1	1732	11 Q54874	Q54874 rattus norv
22	73	12.9	499	10 Q94CT1	Q94CT1 oryza sativ
23	73	12.9	1021	4 Q96M86	Q96M86 homo sapien
24	72.5	12.8	216	3 Q74826	Q74826 schizosacch
25	72.5	12.8	274	16 Q51534	Q51534 pseudomonas
26	72.5	12.8	804	5 Q9NJA3	Q9NJA3 aedes aegyp
27	72	12.7	299	2 Q93D93	Q93D93 streptococc
28	71.5	12.6	284	2 Q9KZP7	Q9KZP7 streptomyce
29	71.5	12.6	1160	5 Q9VJH2	Q9VJH2 drosophila
30	71.5	12.6	1285	10 Q9ZRO1	Q9ZRO1 arabidopsis
31	71	12.5	296	16 Q98CQ4	Q98CQ4 rhizobium 1
32	71	12.5	386	17 Q973P4	Q973P4 sulfolobus
33	71	12.5	412	5 Q9V8D0	Q9V8D0 drosophila
34	71	12.5	995	10 Q22974	Q22974 arabidopsis
35	71	12.5	1295	5 Q23257	Q23257 caenorhabdi
36	70.5	12.5	814	2 Q5767	Q5767 shewanella
37	70.5	12.5	925	16 Q31138	Q31138 pseudomonas
38	70	12.4	414	5 Q19088	Q19088 caenorhabdi
39	70	12.4	437	16 Q927Y9	Q927Y9 listeria in
40	70	12.4	480	2 Q9LAX3	Q9LAX3 streptococc
41	70	12.4	645	16 Q985Q4	Q985Q4 rhizobium 1
42	70	12.4	1030	10 Q9AWS4	Q9AWS4 oryza sativ
43	70	12.4	1038	10 Q94CN3	Q94CN3 oryza sativ
44	69.5	12.3	332	16 Q9HZ43	Q9HZ43 pseudomonas
45	69.5	12.3	386	10 Q49617	Q49617 arabidopsis

## ALIGNMENTS

RESULT 1  
Q90Y41 PRELIMINARY; PRT; 114 AA.  
ID Q90Y41  
AC Q90Y41  
DT 01-DEC-2001 (TREMUREL: 19, Created)  
DT 01-DEC-2001 (TREMUREL: 19, Last sequence update)  
DE 01-DEC-2001 (TREMUREL: 19, Last annotation update)  
DE PREPROSOMATOSTATIN.  
OS Gnathoniemus petersi.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;  
OC Osteoglossiformes; Mormyridae; Gnathonemus.  
OX NCBI\_TaxID=42645;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;  
RT "Characterization of variant somatostatin cDNAs from several  
RT osteoglossomorphs: molecular identification and comparative  
RT analysis".  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF226262; AAK97069.1;  
SQ SEQUENCE 114 AA; 12494 MW; 454DA57A309CA8F2 CRC64;

Query Match 80.6%; Score 456; DB 13; Length 114;  
Best Local Similarity 78.9%; Pred. No. 8.1e-37;  
Matches 90; Conservative 11; Mismatches 13; Indels 0; Gaps 0;  
QY 1 MLSTRVQCALALLSLALAISSVSAAPSADKRLQRLQSLRMAPAGKQELARNTLVLLSEL 60  
Db 1 MLSTRVQCALALLSLALAISSVSAAPSADKRLQRLQSLRMAPAGKQELARNTLVLLSEL 60  
QY 61 AHVNEATELDDMSHGVEQDVLELERAPGPVLAIPRERKAGCKNFYKFTTSC 114  
Db 61 VRVNEALEPDDLSEAGDQEVLELERAGPALAPRERKAGCKNFYKFTTSC 114  
RESULT 2  
Q90Y40

ID	Q90Y40	PRELIMINARY;	PRT;	114 AA.
AC	Q90Y40:			
AC	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DT	PREPROSOMATOSTATIN.			
DE	Chitala chitla.			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu-			
OC	Actinopterygii; Neopterygii; Teleostei; Osteoglossomor-			
OC	Osteoglossiformes; Notopteridae; Chitla.			
OX	NCBI_TaxID=112163;			
OX	SEQUENCE FROM N.A.			
RP	AL-Mahrouki A.A., Irwin D.M., Youson J.H.;			
RA	Characterization of variant somatostatin cDNAs from s-			
RT	osteoglossomorphs: molecular identification and compar-			
RT	analysis.";			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ database:			
RL	EMBL; AF292653; AAK97070.1; -			
DR	SEQUENCE 114 AA; 12561 MW; 4E3C32F58E34F971 CRC64;			
-SQ				
Query Match: 80.2%; Score 454; DB 13; Length				
Best Local Similarity 79.8%; Pred. NO. 1,3e-36;				
Matches 91; Conservative 9; Mismatches 14; Indel				
QY	1	MLSTRVQCALALLSLAIAISVSVAASDAKLQQLQRLQSLMAPAGKQELA		
Db	1	MLSTRIQCALALLSLALPVSVAASDAKLQQLQRLQSLIAPASKQELA		
QY	61	AHVENAEIELDDMHSVGEQDVLELERAPGPVLAAPRERKAGCKNFYWK		
Db	61	AQVENEALSDLSRGADQEVRLERAGPPLAPRERKAGCKNFYWK		
RESULT 3				
ID	Q90XEL	PRELIMINARY;	PRT;	116 AA.
AC	Q90XEL:			
AC	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	SOMATOSTATIN.			
OS	Acipenser transmontanus (White sturgeon).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu-			
OC	Actinopterygii; Chondrostei; Acipenseriformes; Acipen-			
OC	Acipenserinae; Acipenser.			
OX	NCBI_TaxID=7904;			
OX	SEQUENCE FROM N.A.			
RP	TRABUCCHI M., TOSTIVINT H., LIHRMANN I., SOLLARS C., V-			
RA	Dores R.M., Vaudry H.;			
RA	"Polygenic expression of somatostatin in the sturgeon			
RT	transmontanus: molecular cloning and distribution of t			
RT	encoding two somatostatin precursors.";			
RL	J. Comp. Neurol. 0:0-0(2001).			
RL	EMBL; AF395849; AAL13248.1; -			
DR	SEQUENCE 116 AA; 12616 MW; 72E0C3FF6C80650F CRC64;			
-SQ				

```

4
RESULT Q90V42 PRELIMINARY; PRT; 114 AA.
ID Q90V42;
AC Q90V42;
DT 01-DEC-2001 (TReMBEl.re.l 19, Created)
DI 01-DEC-2001 (TReMBEl.re.l 19, Last sequence update)
DT 01-DEC-2001 (TReMBEl.re.l 19, Last annotation update)
DE PREROSOMATOSTRATIN;
DS Pantodon buchholzi (Butterflyfish);
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TAXID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RR "Characterization of variant somatostatin cDNAs from several
RA Al-Mehrhouki A.A., Irwin D.M., Youson J.H.;
RT osteoglossomorphs: molecular identification and comparative
RT analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RD EMBL; AF292651; AAK97068.1;
RS SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 36.7%; Score 208; DB 13; Length 114;
Best Local Similarity 48.7%; Pred. No. 8.5e-13;
Matches 55; Conservative 14; Mismatches 38; Indels 6; Gaps 0

Qy 5 RVOCALALLSLAI-SVSVAAPSDAKURLQLLRSLMAPAGKQELARNTIVELLSELIAHV
Db 5 QVHCILALLGLVLGCGSSSATQLDSRYRSLVQRARAASMGPDWGKLS-VEDLSLIAAA
Qy 64 ENAEIELDKMSHGVEQEDVDLELERA--PGPVLPAPRKACKGNFFWKFTSC 114
Db 64 EAD-MPFGDWSAAEESEGAHLDRSVBPNGV--PPRKACKGNFYWKGFTSC 114

```

```

RESULT      5
Q90143      PRELIMINARY;          PRT;    115 AA.
AC Q90143      2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PREPROSOMATOSTATIN.
OS Osteoglossum bicirrhosum (silver arawana).
OC Euteleostomi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
OX NCBI_TaxId=109271;
RN [1]
RN SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJF databases.
RD EMBL; AF292650; AAK97067.1;
SQ SEQUENCE 115 AA; 12791 MW;  D65FBD7C6F1E4E4D CRC64;

Query Match      35.6%; Score 201.5; DB 13; Length 115;
Best Local Similarity 43.8%; Pred. No. 3,7e-12;
Matches 49; Conservative 17; Mismatches 43; Indels 3; G

Qy 5 RVOCALALLSLAIAISSVAAPS-DAKLRQLQLQSLMAPAGKQELARNTIVELLSELAHV
Db 5 QIHCTVLVLGLVLGLYCPSAASQPDLYRYSFLQRAHAHAANSFPQWSKQAVDELLSLRA

Qy 64 ENAEIELDMSHGVEQEDVDLELERA-PGPVLAPRRKAGCKNFWKFTTSC 114
Db 65 OGF-VPOGAVSRADEEDYVDLERSLELNLPPIRRKAGCKNFWKGFSTC 115

```

```

RESULT 6
Q90Y39 ID Q90Y39 PRELIMINARY; PRT; 120 AA.
AC Q90Y39;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE PREPROSOMATOSTATIN.
DE DE
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Molecular cloning and characterization of white sucker
preprosomatostatin.", to the EMBL/GenBank/DBJ databases.
L Submitted (AUG-2000)
DR EMBL; AF292634; AAK97071.1;
SQ SEQUENCE 120 AA; 13783 MW; 00828D35263E8805 CRC64;

Query Match 34.6%; Score 196; DB 13; Length 120;
Best Local Similarity 43.3%; Pred. No. 1.3e-11;
Matches 52; Conservative 15; Mismatches 37; Indels 16; Gaps 4;

QY 6 VOCALALLSLALATSSVSA-----APSDAKLROLLQRLSMAPAGKQELARNLTVELLSEL 60
Db 6 LHCYLALLGLSLVLCRGADSQLPDMDFHRELLQRLARAIGLATQDWTCKDIELLSQL 65

QY 61 AVHNEATELDDMSHGV-----QEDVDLELERAPGV--LAPREKAGCKNFFWKTFTSC 114
Db 66 SLPEIARE-----NGVSTGGNDLHLEERSAENTNOLYPRERKAGCKNFFWKTFTSC 120

RESULT 7
Q90XE0 ID Q90XE0 PRELIMINARY; PRT; 111 AA.
AC Q90XE0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE SOMATOSTATIN PRO2.
DE DE
OS Acipenser transmontanus (White sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenserinae; Acipenser.
OX NCBI_TaxID=7904;
RN [1]
RP SEQUENCE FROM N.A.
RA Trabucchi M., Testivint H., Lihrmann I., Sollars C., Vallarino M.,
RT "Polygenic expression of somatostatin in the sturgeon Acipenser
transmontanus: molecular cloning and distribution of the mRNAs
encoding two somatostatin precursors.",
RL J. Comp. Neurol. 0:0-0(2001).
DR EMBL; AF395850; AAL13249.1;
SQ SEQUENCE 111 AA; 12748 MW; 4E27DB90896A9025 CRC64;

Query Match 20.7%; Score 117; DB 13; Length 111;
Best Local Similarity 35.3%; Pred. No. 0.00054;
Matches 41; Conservative 15; Mismatches 46; Indels 12; Gaps 4;

QY 6 VOCALALLSLALATSS---VSAAPSAKLRQLQRLSMAPAGKQELARNLTVELLSEL 62
Db 1 MQLRARLVSLMLVYSLRVAVLPGEERLSVHSNREL-----SKERKEGFKLKLGLLDR 55

QY 63 VENEATELDDMSHGVQEDVDLELERAPGVLPAPREKAGCKNFFWKTFTSC 114

```

```

Db 56 VDSVVVLGSDVSPMDLEPDLRLSRLEAIRYNRLSQLPLRARKAPCKNFFWKTFTSC 111

RESULT 8
Q9DDE4 ID Q9DDE4 PRELIMINARY; PRT; 107 AA.
AC Q9DDE4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE SOMATOSTATIN.
DE DE
OS SMST OR SOM.
GN SMST OR SOM.
OC Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-PANCREAS;
RC MEDLINE-99425190; PubMed-10495291;
RX Argenton F., Zecchin E., Bortolussi M.;
RT "Early appearance of pancreatic hormone-expressing cells in the
zebrafish embryo.",
RL Mech. Dev. 87:217-221(1999).
DR EMBL; AJ238017; CAC20110.1;
DR ZFIN; ZDB-GENE-010219-2; smst.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 107 AA; 11839 MW; E12C23E56642EFB CRC64;

Query Match 19.0%; Score 107.5; DB 13; Length 107;
Best Local Similarity 31.6%; Pred. No. 0.0043;
Matches 36; Conservative 17; Mismatches 54; Indels 7; Gaps 3;

QY 1 MLSTRVOCALALLSLALATSSVSAAPSDAKLROLLQRLSMAPAGKQELARNLTVELLSEL 60
Db 1 MASSQLHTATLLCLAM-MAGIICGRSHMLNLSALQASRGTSADIEIPRYSLSLEWLL 59

QY 61 AVHNEATELDDMSHGVQEDVDLELERAPGVLPAPREKAGCKNFFWKTFTSC 114
Db 60 SNSDPAVFQPDSSSLGSLHS--GLELMRD---TKERTGCKNFFWKSRTAC 107

RESULT 9
Q9PRN9 ID Q9PRN9 PRELIMINARY; PRT; 28 AA.
AC Q9PRN9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE GSS-28-SOMATOSTATIN-LIKE PEPTIDE.
DE DE
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE.
RA MEDLINE-96051491; PubMed-8536941;
RA Uesaka T., Yano K., Yamasaki M., Ando M.;
RT "Somatostatin-, vasoactive intestinal peptide-, and granulin-like
peptides isolated from intestinal extracts of goldfish, Carassius
auratus.",
RL Gen. Comp. Endocrinol. 99:298-306(1995).
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 28 AA; 3204 MW; 15D271F677C945BE CRC64;

Query Match 17.5%; Score 99; DB 13; Length 28;

```



Best Local Similarity 81.0%; Pred. No. 0.006; 1; Mismatches 3; Indels 0; Gaps 0;  
Matches 17; Conservative

QY 94 LAPRERKAGCKNFFWKFTSC 114  
DB 8 LPARERKAGCKNFFWKFTSC 28

RESULT 10

QYPRVO PRELIMINARY; PRT; 25 AA.  
AC Q9PRVO; 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE SOMATOSTATIN-RELATED PEPTIDE  
OS Anguilla japonica (Japanese eel)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
OC Anguillidae; Anguilla.  
NCBI\_TaxID=7937;  
N [1]  
RP SEQUENCE.  
RX MEDLINE=95053622; PubMed=7525832;  
RA Usaka T., Yano K., Yanasaki M., Nagashima K., Ando M.;  
RT "Somatostatin-related peptides isolated from the eel gut: effects on  
ion and water absorption across the intestine of the seawater eel.";  
RL J. Exp. Biol. 188:205-216(1994).  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.  
SQ SEQUENCE 25 AA; 2860 MW; BFC672143A04A3F5 CRC64;

Query Match 15.4%; Score 87; DB 13; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.076;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 97 RERKAGCKNFFWKFTSC 114  
DB 8 RERKAGCKNFFWKFTSC 25

RESULT 11

QYRIP8 PRELIMINARY; PRT; 105 AA.  
AC Q9RIP8; 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE PREPROCORTISTATIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
N [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=99453346; PubMed=10521599;  
RA Calbet M., Guadano-Perraz A., Spier A.D., Maj M., Sutcliffe J.G.;  
RA Przewlocki R., de Lecea L.;  
RT "Cortistatin and somatostatin mRNAs are differentially regulated in  
response to kainate.";  
RL Brain Res. Mol. Brain Res. 72:55-64(1999).  
N [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RA Calbet-Murro M., de Lecea L.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050136; AAD51127.1;  
DR MGI; MGI:109538; Cort.  
DR InterPro; IPR004250; Somatostatin.  
DR Pfam; PF03002; Somatostatin; 1.

FT CHAIN 88 105 CORTISTATIN.  
SQ SEQUENCE 105 AA; 11061 MW; FODCD40052AB8E95 CRC64;

Query Match 15.0%; Score 85; DB 11; Length 105;  
Best Local Similarity 53.6%; Pred. No. 0.64;  
Matches 15; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 87 ERAPGPVLAPRERKAGCKNFFWKFTSC 114  
DB 77 QERPPQPPHLDKKPKCNFFWKFTSC 104

RESULT 12

O63803 PRELIMINARY; PRT; 715 AA.  
AC O63803; 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)  
DE 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE XLALPHAS PROTEIN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
N [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95089824; PubMed=7997272;  
RA Kehlenbach R.H., Matthey J., Huttner W.B.;  
RT "Xlas is a new type of G protein.";  
RL Nature 372:804-809(1994).  
N [2]  
RP SEQUENCE FROM N.A.  
RA Klemke M., Kehlenbach R.H., Huttner W.B.;  
RT "Two overlapping reading frames in a single exon encode interacting  
proteins - a novel way of gene usage.";  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
N [3]  
RP SEQUENCE OF 1-367 FROM N.A.  
RC STRAIN=WISTAR;  
RA Wang Y.Z., Kehlenbach R.H., Huttner W.B.;  
RT "The XL-domain of rat Xlas is encoded by a single exon.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X84047; CAC39211.1;  
DR EMBL; AF093569; AAD03032.1;  
DR HSP; P04896; IAZT.  
DR InterPro; IPR001019; Gprotein\_alpha.  
DR Pfam; PF00503; G-alpha; 1.  
DR PRINTS; PR00318; GPROTEINA.  
DR SMART; SM00275; G-alpha; 1.  
SQ SEQUENCE 715 AA; 78232 MW; 68F9AAAC18159896 CRC64;

Query Match 14.0%; Score 79.5; DB 11; Length 715;  
Best Local Similarity 30.1%; Pred. No. 20;  
Matches 40; Conservative 13; Mismatches 55; Indels 25; Gaps 7;

QY 2 LSTRVOCALALSLALSLAIS---SVSAAPSADKLK-----QLQSLMAPA-----GK 45  
DB 171 LPARAAAAAARAARAAACRSASAPRAHLRPPSPPEIQVADPTTPRPAFSAFDPK 230  
QY 46 QELARNTL-VELLSELAHVENAELELDDMSHGVEQEDVLELERAPGVPLAPRERKAG-- 102  
DB 231 YERGSCCRYEAAAGICEISSDESEGATGCFQ--WLLRRNRPG---QPRSHTVGSN 285

QY 103 -CKNFFWKFTSC 114  
DB 286 PVRNFFARAFGSC 298

RESULT 13  
O01392



